A Search for Transiting Extrasolar Planets and Variable Stars in the Galactic Plane

---

A thesis submitted in partial fulfilment of the

requirements for the Degree of Doctor of Philosophy in Astronomy

in the University of Canterbury

by Veronica R. Miller

Department of Physics and Astronomy

University of Canterbury

2009

Abstract

This work describes the observations and results found from a photometric survey of a 0.5° by 0.5° area of the Galactic Plane performed using the 2.2 metre ESO telescope at La Silla, Chile. The dataset comprises a total of 267 images with 204 from a 16 day observation run in 2005 and 63 from a six week observation run in 2002. The new image subtraction reduction algorithm implemented on this data resulted in more than 500,000 lightcurves with a magnitude limit of R $\sim$ 24.5. The precision of the data following reduction is suitable for transit searches as well as identification of variable stars.

Resulting from the transit search was an initial list of 31 candidates, reducing to 23 on further examination. Nine candidates were eliminated by examination of the images and the remaining list re-reduced. After this reduction three good candidates remain. These candidates have periods from 1.2840 to 2.6269 days and depths of around 75 mmags. These three candidates require followup of either multi-colour photometry or spectroscopy to determine their nature.

The variable star search resulted in detections of 1475 variable stars of different types. The largest portion were eclipsing binary stars. A number of the contact binaries have possible low mass-ratios and there are also a number of contact and detached binaries which may contain low-mass components. Three of the contact binaries were found to have periods at the known period cut off including two with periods lower than any previously published. Also identified were two possible pre-main sequence detached eclipsing binaries. The binary fraction of the field was calculated from the observed contact binaries to be 46% $\pm$ 4%. There are a number of possibilities for further data mining of the survey.

Contents

List of Figures

List of Tables

Acknowledgments

The work of a PhD includes input and support from many people.

First my thanks go to my supervisor, Dr Michael Albrow, whose guidance and advice have made this possible.

I would next like to acknowledge the help given by Dr Jean-Philippe Beaulieu and Dr Cristina Afonso in particular for their assistance in obtaining further observations.

I also wish to acknowledge the support and help given by other members of the department. Particular thanks go to Dr Orlon Petterson for computer support and assistance, Dr Mike Reid for supervisory comments and Dr Adrian MacDonald for keeping me on track.

Much of the work off this research would not have been possible without the use of the University of Canterbury Super Computer Bluefern. My grateful thanks go to the UCSC team with particular thanks to Tony Dale.

During the course of this research I was supported by a Physics and Astronomy Departmental PhD Scholarship as well as a large amount of support from the Marsden Fund through my supervisor. My trip to the Transiting Extrasolar Planets Workshop and subsequent stay was supported by a scholarship from the Max-Planck Institut fuer Astronomie. Attendance at the IAU symposium 253 on Transiting Planets in Boston was made possible by a travel grant from the IAU.

I would like to thank the graduate students of the department (both past and present) for assistance in various matters. In particular I wish to acknowledge my office mates for answering strange questions at all hours of the day.

Almost finally I wish to thank my family and friends for all their support throughout.

And lastly I would like to thank my husband for all his support and love. Thanks are not really enough but I expect you know how I feel.

Chapter 1

Introduction

1.1 History of extrasolar planet observations

Extrasolar planets do not have a long history of detection though suppositions have been made about the possibility since it was determined that the sun was a star. In the 1940's three papers on the possibilities of unseen companions to stars of varying properties were published: van de Kamp (1943, 1944) and Strand (1944). Strand discovered several planet-like companions in this astrometric work. The confirmation of the existence of one of the planet-like companions in 1951 led to a short article by Struve (1952) discussing planet-like companions and the possibilities of detecting them. While these 'unseen companions' were later found to be stellar in nature the methods used in the research can be applied to the search for planets. Strand and others used direct photography on the systems they were investigating. The size of the planet-like companions discovered, in general $> 0.5M_{\odot}$, were such that this was found to be a suitable method, but the limits of the photographic methods at that time gave no possibility of detection of objects like Jupiter or even those ten times the size at any real distance from the host star. The knowledge that stellar companions can exist very close to each other allowed the point to be made that it is indeed possible that a planet might actually exist at such a distance to leave it with

a period of around one day. For instance, a $10M_J$ object would have an effect on the radial velocity of its host star such that it would be measurable with spectrographs available at the time of Struve's paper. The further implication of a planet of this size and distance is that it would, in the optimistic case, give a transit of depth 0.02 mag, a variation ascertainable by the photoelectric efforts of the day.

In spite of Struve's optimistic paper, no planetary companions were confirmed for several decades. The probability that any single star observed is hosting a planet is relatively low and it was not in fact until 1992 that the first confirmed extrasolar planet was found. This detection was made by observations of variations in the timing of a pulsar (Wolszczan and Frail, 1992). While this discovery was a wonderful confirmation of extrasolar planetary existence, astronomers hoped to discover a planet around a star more like our own and efforts to this end continued. In 1995 the first planet orbiting a sunlike star was discovered by the radial velocity method, (Mayor and Queloz, 1995) and was followed shortly by several others also detected by the radial velocity method, (see Marcy and Butler (1996),Butler and Marcy (1996) and Cochran et al. (1996)). Over the last decade the number of extrasolar planets found each year has grown by leaps and bounds - in the last year alone more than 60 planets have been confirmed, bringing the total to more than 300 confirmed planets, with a number of unconfirmed planets.

1.2 Definition of an extrasolar planet

For many years the words 'extrasolar planet' were used but not defined and indeed even the definition of what comprises a planet in our own Solar-System was in doubt. After much debate the International Astronomical Union voted on the definition of a Solar-System planet at the General Assembly in 2006:

A "planet" is defined as a celestial body that (a) is in orbit around the Sun, (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a

hydrostatic equilibrium (nearly round) shape, and (c) has cleared the neighbourhood around its orbit.

In 1999 the Working Group on Extrasolar Planets (WGESP) was created by the IAU with the purpose of acting as focus for extrasolar planetary research. In February 2003 the WGESP made a statement regarding the working definition of an extrasolar planet. This definition is described as being applicable for all those cases in which there are already confirmed detections and will not necessarily hold for all future definitions. The definitions made here will undoubtedly evolve as more information about extrasolar planets comes available. The definition herein also depends on the definition of Solar System planets.

1) Objects with true masses below the limiting mass for thermonuclear fusion of deuterium (currently calculated to be 13 Jupiter masses for objects of solar metallicity) that orbit stars or stellar remnants are planets (no matter how they formed). The minimum mass/size required for an extrasolar object to be considered a planet should be the same as that used in our Solar System.

2) Substellar objects with true masses above the limiting mass for thermonuclear fusion of deuterium are brown dwarfs, no matter how they formed nor where they are located.

3) Free-floating objects in young star clusters with masses below the limiting mass for thermonuclear fusion of deuterium are not planets, but are sub-brown dwarfs (or whatever name is most appropriate).

These statements do not depend in any way upon the formation method of the object under consideration, as such different objects of the same size and possibly formation may in fact be defined with different names solely due to their location.

1.3 Motivation for extrasolar planet research

The Solar System has, for many years, been regarded as a model for planetary systems through the universe, though with an extremely limited sample size. Knowledge of the Solar System leads to the thought that all planetary systems will consist of small rocky planets close to the star and large gaseous planets at a greater distance. A lack of data of other planetary systems has limited the research that can be performed on planetary systems without assuming that the design of our Solar System applies in the wider Universe. As such, theories have depended on the Solar System being a model for all planetary systems. Due to the fact that the most numerically successful techniques of planet detection (radial velocity, §1.4.1 and transits, §1.4.7) are blind to the presence of Solar System analogues, only one system has been detected which appears to follow any of the same expectations as the Solar System, (Gaudi et al., 2008). The detection of a large number of planets which do not follow the lines of the solar system has caused a large-scale restructuring of theories of planet formation, evolution and distribution has occurred. Further observations and detections of extrasolar planets are required to answer such questions as:

- whether planets exist around binary stars or in clusters
- how planetary systems evolve
- whether there is a dependence on stellar metallicity or mass
- if there are other planetary systems like the Solar System in the Universe

Characterization of a large sample of extrasolar planets will allow theoretical predictions to be tested and improved. Detection of as many different possible planet arrangements and types will allow theories to be probed, tested and reworked to fit the Universe as a whole. Determination of what conditions are required for the formation of planets will answer whether planets are able to form in multiple star

systems or in closely packed clusters and indeed if there any special requirements of the host star for the formation of planets.

Another area of interest, in particular to the public at large, is the search for extrasolar Earth-like planets, in particular the detection of such planets in the habitable zone. A planet in the habitable zone is at a distance from the star at which the surface temperature of the planet would allow liquid water.

Observations of extrasolar planets allow for a greater understanding of our place in the Universe and how the Solar System came to be.

1.4 Extrasolar planet search methods

There are a number of different search methods for extrasolar planets, each with a different style of observing and providing different information. Methods include those using spectroscopic observations (radial velocity) and photometric observations (astrometry, transits, microlensing, timing, direct imaging and circumstellar disks). The detections made by each method are tabulated in Table 1.1.

Search Method	Planetary Systems	Planets
Radial Velocity	263	307
Transit	54	54
Microlensing	7	8
Direct Imaging	9	11
Timing	4	7
Astrometry	0	1
Circumstellar Disks	0	1

Table 1.1: **Planet detections:** Tabulation of the number of planet detections by search method. In the case of astrometry the only detection made was of a previously discovered planet and in the circumstellar disk and timing methods the detections have been made in previously known systems. Numbers taken from the Extrasolar Planets Encyclopedia (Schneider et al., 2008).

Ideally an extrasolar planet would be detected by every method giving the maximum amount of data on that planet. Due to differing observational and physical requirements for the different methods this is not possible, but it is often possible to make observations of the same planetary system with more than one method. For instance, the radial velocity technique is commonly used as a confirmation of transiting planets through the determination of the planet-star mass ratio.

1.4.1 Radial-velocity method

The most productive technique thus far in detecting extrasolar planets is the radial velocity technique. This detects Doppler shifts in the radial velocity of stars and was first implemented as a planet finding technique by Marcy and Butler (1993). A planet orbiting a distant star causes the star to move around the common centre of mass in a small circular orbit in response to the planet's gravity. The movement towards or away from the observer will cause a small shift in the observed spectrum of the star. This shift can be detected by observing lines in the stars spectrum and comparing them to known laboratory wavelengths. As the planet moves around its host star the spectrum of the star will shift back and forth in wavelength.

Despite the extrasolar planet's gravity being small in comparison to the star's, providing the planet is massive enough and close enough to the star, the change in velocity will be large enough to be detectable with current spectroscopic instrumentation (i.e. $\sim$1 ms^{-1}). Limits in spectroscopes mean that only small numbers of stars can be observed at one time, and only the brighter and closer stars give sufficient signal to noise. The radial velocity method allows the period and minimum mass of a planet to be determined. Without knowing the inclination of the orbit it is not possible to determine actual mass, but if we can detect a transit (§1.4.7) of the planet or use the radial-velocity method to confirm a known transit the inclination will be known and the actual mass calculated.

The radial velocity detection technique has thus far been responsible for ap-

proximately 80% of extrasolar planets detections and includes several multi-planet systems. Several different spectrometers are working on radial-velocity surveys in order to detect extrasolar planets. The Anglo-Australian Planet Search (AAPS) is being run on the 3.9 m Anglo-Australian telescope in Australia and is searching nearby Solar type stars. The Geneva Observatory runs three search programs using spectrographs: HARPS (High Accuracy Radial velocity Planetary Search (Pepe et al., 2000)) at the European Southern Observatory (ESO) 3.6m telescope at La Silla, Chile; *Elodie* (Queloz et al., 1998) at the 1.93m telescope at L'Observatoire d'Haute Provence, France; and *Coralie* (Queloz et al., 2000) at the 1.2m Swiss telescope at La Silla, Chile. The California and Carnegie Planet Search runs programs at the Keck Observatory in Hawaii using HiRES (High Resolution Échelle Spectrograph (Sozzetti et al., 2005)) and using previous observations from the Lick Observatory in California (Fischer et al., 1998). HARPS has been particularly successful both in finding extrasolar planets by radial-velocity and confirming extrasolar planets found by transits.

1.4.2 Microlensing method

A gravitational lens occurs when the light from a distant bright object is bent by the gravity of a massive foreground object - often forming two or more images. Albert Einstein, while not the first to suggest the idea, is most often credited with the notion and the characteristic angular scale, the Einstein Ring radius is named for him.

Microlensing is the effect of gravitational lensing on a smaller scale. A single star (the lens) can bend the light of a background star (the source) forming two unresolved images and increasing the amount of light visible from Earth over a timescale of days or weeks depending on the lens-source proper motion. If the star acting as the lens (the foreground star) has a planet in orbit, the planet may also have gravity enough to act like a smaller lens. The result of a extrasolar planet

involved in the lens causes a spike in the smooth curve of the lightcurve of the lensed star, (Figure 1.1). Several surveys are involved in the observation of different parts of the sky to detect the initial increase in light the indicates a microlensing event, (see OGLE (Udalski et al., 1992),MOA (Bond et al., 2002) for more). The star is then followed up with more frequent observations as its nature becomes clear. The microlensing events typically increase in brightness over a number of weeks with increased monitoring occurring as the brightness peaks. Several followup groups (including the Planetary Lensing and Anomalies Network (PLANET) collaboration from whom the original data used in this research was obtained) use alerts from the survey teams to perform large numbers of observations of the events. Microlensing has found a number of planets so far including the most earth-like planet to date, (Beaulieu et al., 2006).

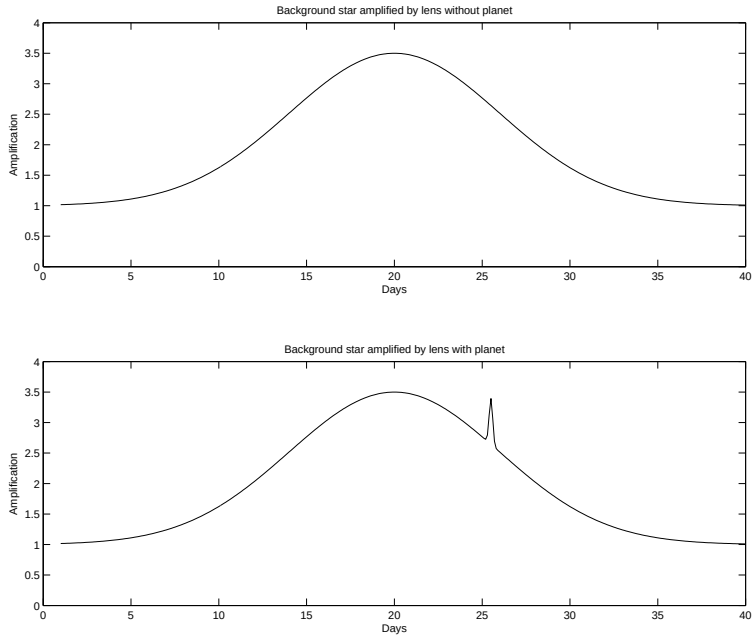


Figure 1.1: **Theoretical microlensing lightcurves** Upper panel shows a typical lightcurve of a galactic microlensing event. The lower panel shows the same event, but with a planet located near to the lens star.

1.4.3 Astrometric method

Astrometry is the oldest extrasolar planet search method and uses measurements of the position of a star in the sky. The gravity of an extrasolar planet orbiting the star has an effect on the position of the star and causes the planet's motion to scribe a small circular or elliptical orbit. This works in a similar way to the effects measured by radial velocity variations though the alignment of observer and planetary orbit are different. At this time, measurements taken from earth are not precise enough to detect the small changes caused by a planet. The Hubble Space Telescope was however able to take astrometric measurements of a previously known planet Gleise 876, (Benedict et al., 2002) and it is hoped that future space missions will discover many more planets using this method. The astrometric method is most sensitive to planets in large orbits and so is complementary to the selection effects of other search methods. The long orbits will also require long observation times and it may takes years or decades of careful measurements for detections to be made.

1.4.4 Pulsar timing method

A pulsar is a neutron star, a small, very dense star which emits a large beam of radio waves. As the star rotates, the beam of radio waves passes by Earth at extremely regular intervals. A normal pulsar will have an extremely regular frequency and small anomalies in the timing of these can be used to track the movement of the star. A planet will move the pulsar in a small circle as detected in radial velocity and astrometry measurements (§1.4.1 and §1.4.3) and thus will affect the timing of the pulse as it reaches earth. Observations of pulsars were not being searched for extrasolar planets, but several planets have been discovered around pulsars. This includes the first ever detected extrasolar planet in 1992, (Wolszczan and Frail, 1992). Pulsars in general are relatively rare and so there are not expected to be large numbers of planets found in this way though the high precision possible for

measurements of the frequency of a pulsar makes a planet detection easy.

1.4.5 Direct imaging method

While in general the light coming from a planet is too faint to detect as separate from the host star, in some cases it is possible to directly image extrasolar planets. There are several projects underway to equip larger telescopes with extrasolar planet finding instruments. The Gemini telescope (NIRI + ALTAIR), VLT (NAOS-CONICA) and Subaru (Suprime-Cam and Hyper Suprime-Cam) all have such instruments either just in use or soon to be in use. Planets most likely to be found will be massive, with large orbits around small host stars. In 2004 observers at the VLT imaged a companion to the brown dwarf 2M1207, (Chauvin et al., 2004) using the NACO instrument and its status as an extrasolar planet was confirmed at the end of 2005. The planet is believed to have an orbit of more than 40 AU with a mass of several times Jupiter's. Following this discovery several other possible planets around dim stars have been imaged but their planetary status has not yet been confirmed and it is possible that they are instead small brown dwarfs. Most recently a planet around a solar type star was imaged by the Gemini North telescope using the NIRI camera and ALTAIR adaptive optics. The planet is approximately $8M_J$ and at a distance of 330 AU,(Lafrenière et al., 2008).

1.4.6 Extrasolar planets in circumstellar disks

Planets may also be detected by perturbations of dust disks around stars. Dust forms disks around many stars and is detected because it absorbs light and re-radiates it at infrared wavelengths. While the mass of the disks is in general far less than the mass of earth, the large surface area the dust contains allows enough light re-radiated that the disks often outshine the host star in the infrared. Both the Hubble and SPITZER space telescopes have instruments capable of imaging the

infrared light from circumstellar disks. Features in the disks can be observed and large ones may indicate the presence of full-sized planets along with the asteroids which are far more common. ϵ Eridani has such features in its disk and indicate the presence of a secondary planet at 40 AU as well as the already known inner planet (Greaves et al., 2005). Most recently a planet at 8 AU was imaged inside the disk of β Pictoris using the VLT (Lagrange et al., 2009). The shape of the trails in the disk and the tracks of comets falling into the star suggest the mass of a giant planet.

1.4.7 Transit method

The second most productive method for finding extrasolar planets is the transit or dimming method. Providing the plane of the planet's orbit is in line of sight to the observer, a planet transiting its host star will cause the light coming from the star to dim, Figure 1.2. The search for this change in the light from the star can performed on large numbers of stars at once providing the precision of the lightcurve is suitable. A transit of a planet like Jupiter orbiting a star like our Sun will produce a change in flux of around 1%. Other planets, such as those found by the radial velocity method, are often somewhat larger than Jupiter as well as in a smaller orbit around the host star. As a consequence the transit causes a larger change in the lightcurve and the effect can be up to 5%. The number of planets found thus far by the transit method comprises approximately 13% of the total planet detections, with the percentage of planets detected by the transit method increasing all the time.

The transit method is most sensitive to massive planets orbiting close to the host star, much like the radial velocity method. As such it is most likely to detect those planets known as hot Jupiters. Hot Jupiters are extrasolar planets of masses

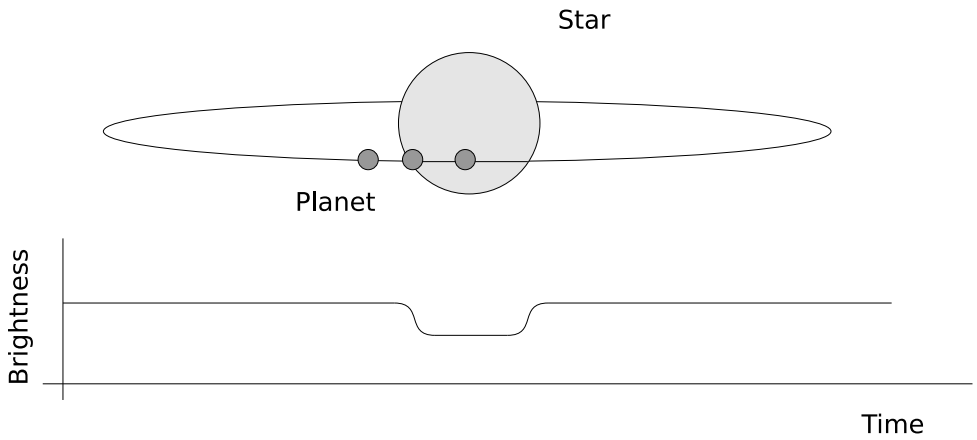


Figure 1.2: **Theoretical transit lightcurve** The lightcurve here shows the change in light levels as they correspond to the position of the planet as the planet transits its host star.

similar in size to that of Jupiter that orbit very close in to the host star. Typical periods of a hot Jupiter are between 3 and 5 days. The transiting extrasolar planets detected thus far, whether by followup of a radial velocity detection or discovered originally by the transit method, are all of the hot Jupiter type. A number of these have also been denoted "very hot Jupiters" - these are even closer to their host star, with periods of less than 2 days.

A visible transit requires that the plane of the planet's orbit be in the line of sight to the observer. The mass of a planet can be determined if it is observed photometrically through a transit and subsequently with spectroscopy. The number of stars with suitable transiting planets (§1.4.7.3) is small in relation to the number of observable stars, so in order to find transit candidates it is necessary to observe large numbers of stars at once. For transiting planets this can be accomplished with a wide field camera for a crowded star field (for example Pan-Planets, OGLE, MOA, this research), or, by rapidly sampling sparse fields of bright stars with a smaller camera (for example SuperWASP, KELT, XO) (see §1.4.7.2 for details on surveys.).

1.4.7.1 Advantages and disadvantages of the transit method

Detecting an individual planet around a star has a far lower probability of succeeding than the radial velocity method because of the geometrical requirements for a transit. To offset this however is the ability to observe thousands of stars in a single field of view and acquire lightcurves of all stars simultaneously. Transit surveys also have the ability to make use of small telescopes, in particular 10 - 20 cm inexpensive commercially available telescopes (see Alonso et al. (2004); Pollacco et al. (2006) for more information). Telescopes of this size can often be entirely dedicated to transit surveys, not subject to the competition involved in obtaining telescope time for use in radial velocity surveys. A major advantage of a transit search is the gain in information from a detected transiting extrasolar planet over any other detection method.

Detection of an extrasolar planetary transit gives measurements of the orbital period, orbital distance, orbital inclination, planetary radius and, in conjunction with radial velocity followup measurements, a true measure of mass and subsequently the density of the planet. The transit method is the only detection method that enables measurements of all these parameters. Transiting planets thus provide important constraints on theories of planetary composition and structure.

Followup photometry of detected transiting extrasolar planets provides a wealth of information:

- additional planets may be detected,
  - the orbits of additional planets in a multiple planet system have a large possibility of being coplanar and thus will also transit. There is a possibility of finding terrestrial planets further out in the system within the habitable zone along with the close in hot Jupiters (Raymond et al., 2004)
  - using long time-scale observations in the detection of additional planets in the system by means of transit timing variations (Agol et al., 2005). There is a possibility to detect additional planets down to Earth-mass by the effects they have on the period of the primary planet.
- high precision photometry has the possibility of detecting moons (Deeg, 2002) and rings (Barnes and Fortney, 2004),
- spectroscopic observations of the transit can be used to detect atmospheres (Charbonneau et al., 2002; Vidal-Madjar et al., 2003, 2004; Marley, 2008),
- observations of infrared emission in secondary eclipses (Charbonneau et al., 2005; Deming et al., 2006) which can also give information on the atmosphere,
- observations of the Rossiter-McLaughlin effect through the transit allows determination of the alignment of the orbital and spin axes and allows the possibility of constraining formation and evolution scenarios (Winn, 2007) and

- a large dataset of high precision lightcurves is obtained from any transit survey and allows for searches to be made for variable stars in particular eclipsing binaries with low-mass companions.

Transit survey observations can be made of any part of the galaxy allowing measurements of very different types of stellar environments. Transit surveys can be tailored to investigate specific areas like globular clusters (for example Weldrake et al. (2005, 2007a); Gilliland et al. (2000)); open clusters (for example von Braun et al. (2005); Mochejska et al. (2002); Hartman et al. (2007));or the Galactic Bulge (for example Gould et al. (2006); Sahu et al. (2006)). Determinations can be made on the dependence of planet formation on stellar mass and metallicity. Transit surveys are not in anyway volume limited unlike radial velocity surveys which are thus far constrained to $\simeq$ 200pc and can therefore obtain new information. A large number of stars can be observed at a single time with a wide-field camera allowing large portions of the sky to be searched in a single survey, other methods require observations of single stars and are far more consuming of telescope time. An additional scientific spinoff: a large dataset of high precision lightcurves is obtained from any transit survey and allows for searches to be made for variable stars, in particular eclipsing binaries with low-mass companions

Most problematic to transit detections is the need for further followup observations to determine the nature of the transiting object. It is possible for planet-like transits to occur with eclipsing binary stars: where the primary star is much larger than the secondary; or with blended stars: there is an undetected object also contributing light to the system (see §5.5 for further descriptions). In what is becoming a large number of cases, there is no easy way for radial-velocity measurements to be taken of the fainter stars observed by wide-field transiting surveys. Transit surveys have a high false positive rate, however the scientific information returned from a single detection justifies the effort put into weeding out false positives.

Transit surveys at this stage have a bias towards detection of large planets

close to their host stars. The current lower limit of detections is $\approx 0.4\ R_J$. The two space missions CoRoT and *Kepler* are aiming for detections of Earth-size transiting extrasolar planets (see §1.4.7.2 for further information).

1.4.7.2 Transit search strategies and surveys

Over the last decade many transit surveys have started observations, including one space-based program. The surveys are composed of two major types; shallow, wide-field surveys and deep, narrow surveys. Both types of surveys involve observing many stars at regular intervals over a period of time.

There are a large number of different transit surveys ongoing including:

- ground-based
  - designed for transit detection (for example SuperWASP, KELT, TrES and XO)
  - designed for other photometric observations (for example HATnet, OGLE and MOA)
- space-based surveys (for example CoRoT and *Kepler*)
- followup observations of known extrasolar planets (for example MOST and EPOXI) to obtain and refine parameters

The Optical Gravitational Lensing Experiment (OGLE), Udalski et al. (2002) observes a large area of the sky as part of its microlensing survey and the observations from the survey are ideal for transit searches. Since 2002 the OGLE collaboration has discovered eight planets via the transit method. A second microlensing survey collaboration, Microlensing Observations in Astrophysics (MOA), has also begun searching their survey results for transiting planets but are yet to identify any. The large amount of data obtained in the search for microlensing events is ideal for transit searches. Both the MOA and OGLE data is collected with single large telescopes

with wide-angle cameras and the observations concentrated in reasonably crowded fields. There are also other smaller surveys specifically designed for transit search working with similar equipment. These observations tend to have a shorter overall time scale.

More recently is the development of several productive surveys using multiple small telescopes. The HATnet project, (Bakos et al., 2004; Hartman et al., 2004; Shporer et al., 2007) has detected nine transiting planets in the last few years. HATnet was originally designed for observing bright variable stars and currently consists of a set of four small robotically controlled telescopes. Plans are currently in the works for a second HATnet installation in the southern hemisphere. Along with HATnet the Harvard-Smithsonian Centre for Astrophysics (CfA) is also the homebase for another survey project. The Trans-Atlantic Exoplanet Survey (TrES) survey, (O'Donovan et al., 2007; Alonso et al., 2004) has had success (four planet detections) with using three small aperture telescopes to search for transits. Also using a small aperture telescope is the Kilodegree Extremely Little Telescope (KELT), (Pepper et al., 2004) a wide-field small aperture transit survey of bright stars and the XO telescope consisting of two 200mm telephoto camera lens working like a pair of binoculars. KELT has not yet found planets but the XO telescope, McCullough et al. (2005) has been successful in detecting five transiting planets (also making use of a large number of amateur observations to fill in the lightcurves of these planets).

Similar to the XO telescope, the SuperWASP project based in the United Kingdom (Cameron et al., 2007) has been extremely successful with confirmations of 15 planets. SuperWASP is comprised of two robotic telescopes located at the South African Astronomical Observatory in South Africa and La Palma in the Canary Islands. The telescopes each consist of eight wide angle telephoto lenses with cameras that monitor the sky in unison for transits. Also based in the UK is the MONITOR survey, Irwin et al. (2007b) which observes young open clusters and uses the data from these observations to also search for transits.

For fainter and more distant stars, transit detections surveys are moving into space with two specifically designed satellites to be up and running with detections by the end of 2009. Space telescopes have the advantage of being outside the Earth's atmosphere and can obtain far more precise photometry even with small telescopes.

The first of these satellites specifically designed for transit searches was launched at the end of 2006. The CoRoT (Convection, Rotation and Planetary Transits) satellite, (Bordé et al., 2003) will observe a single field continuously for four months before moving on to the next field. The data from the satellite will be very high precision (one part in a million) and with very even temporal sampling. The number of transit detections from CoRoT is expected to be high and dramatically increase the total number of known transiting planets. In the time since launch, four planets and one brown dwarf have been detected (Barge et al. (2008), Alonso et al. (2008), Bouchy et al. (2008), Aigrain et al. (2008), Moutou et al. (2008) and Deleuil et al. (2008)). In the testing phase of CoRoT's software, lightcurves from the Berlin Exoplanet Search Telescope (BEST), (Kabath et al., 2007), a 20cm telescope with observations of transit detection quality over a two year period, were used to conduct blind-tests. The BEST program has not detected any transits on its own.

While the Hubble Space Telescope (HST) was not in any way designed for transit searches it has been used successfully (and unsuccessfully) for this purpose. The Sagittarius Window Eclipsing Extrasolar Planet Search (SWEEPS, Sahu et al. (2006)) used the Advanced Camera for Surveys (ACS) to observe a field in the Galactic Bulge for seven continuous days. SWEEPS has found two confirmed transiting extrasolar planets.

Kepler (Basri et al., 2005) is another survey satellite with an expected launch date of early 2009. The *Kepler* satellite is designed to search for earth-like planets. With a mission length of three years and high precision photometry, there are hopes that the survey will find long-period small planets.

There are also groups working on the detection of secondary planets within

known planetary systems and those endeavouring to detect transits where the initial detection was made by a different method (most commonly radial velocity). The MOST satellite is a Canadian small satellite designed for stellar variation observations. The satellite is performing followup photometry on known transits and has hopes to detect secondary planets in these systems by search for transit timing variations caused by the effect of a secondary planets gravity, (Croll et al., 2007). The EPOXI satellite mission (Christiansen et al., 2008a) is also working on detecting secondary planets by completing intense photometric observations of five of the known extrasolar transiting planets.

All possible transit detections need followup observations to confirm the detection. This involves obtaining radial-velocity measurements of the star in question. The large transit surveys often have access to large telescopes in order to obtain these in a timely manner. Smaller surveys and deep surveys have more difficulty obtaining the observations required. There are currently 54 unconfirmed extrasolar planets listed in the international database (Schneider, 2008) and a large number of these have come from transit surveys.

Other methods of filtering out blended eclipsing binaries and other false positives include using the colour of the stars and further individual photometric observations with high cadence. Target of opportunity observations on the SPITZER telescope have been used in some cases to perform followup or secondary observations on known transiting planets such as observations to investigate water and detect secondary eclipses, (Ehrenreich et al., 2007; Deming et al., 2007).

There are two other large, ground based, transit surveys beginning in the next few years. The SkyMAPPER telescope (Keller et al., 2007), which will be operating out of Siding Springs Observatory in Australia, is to perform a wide-field sky survey of the Southern hemisphere in a number of different filters. The data from this survey will be ideal for transit searches and there are a number of people working on it for this purpose. The Pan-Planets transit search (Afonso and Henning, 2007),

based in Hawaii, will use the PANSTARRS 1.8 metre telscope and wide field camera. PANSTARRS (Panoramic Survey Telescope and Rapid Response System) has an immediate goal of detecting near-Earth asteroids and comets, but over its lifetime will produce a huge amount of data, a large portion of which will be suitable for detecting transits.

1.4.7.3 Predictions for survey yields and problems

In a proposal for observing time for transit observations it is important to make accurate predictions of the planetary yield. Here we discuss the likely yields and problems with such.

For a transit to be observed the orbital plane must be nearly edge on. If the inclination (i) is less than 90° the shape of the transits narrow and become shallow, disappearing as $cos\ i\ <\ (R+r)/a$, where R and r are star and planet radii and a the semi-major axis. For a random orientation the probability for the observation of a transit is given by Equation 1.1.

$$P_t = \frac{R+r}{a} \approx 0.1 \left(\frac{P}{4d}\right)^{\frac{2}{3}} \left(\frac{R}{R_\odot}\right) \left(\frac{M}{M_\odot}\right)^{-\frac{1}{3}} \tag{1.1}$$

For a large planet orbiting close to the host star (P ≈ 4 days, a ≈ 0.05 AU) the transit probability is approximately 10%. An Earth-like planet, however, has only 0.05% probability of transit.

To ensure accurate identification, three or more transits must be observed and even this, in a weak detection (depth ≈ lightcurve noise), is not enough to confirm the transit without further followup (see §5.5).

Initial predictions of planetary yield from transit surveys by Horne (2002) and Brown (2003) were based on the assumption of white noise, i.e. Gaussian noise that is uncorrelated between data points. In fact, as determined only a few years later by Pont et al. (2006), noise in high precision time-series photometry has a

significant low frequency component. This noise is not uncorrelated but co-variant, and is known variously as co-variant, systematic or, most commonly, red noise. Figure 1.4.7.3 shows examples of noise in a lightcurve. The top panel shows white noise, the middle panel red noise and at the bottom the combination. Each panel of the figure shows the same dispersion of points with only the frequency of the variations changing.

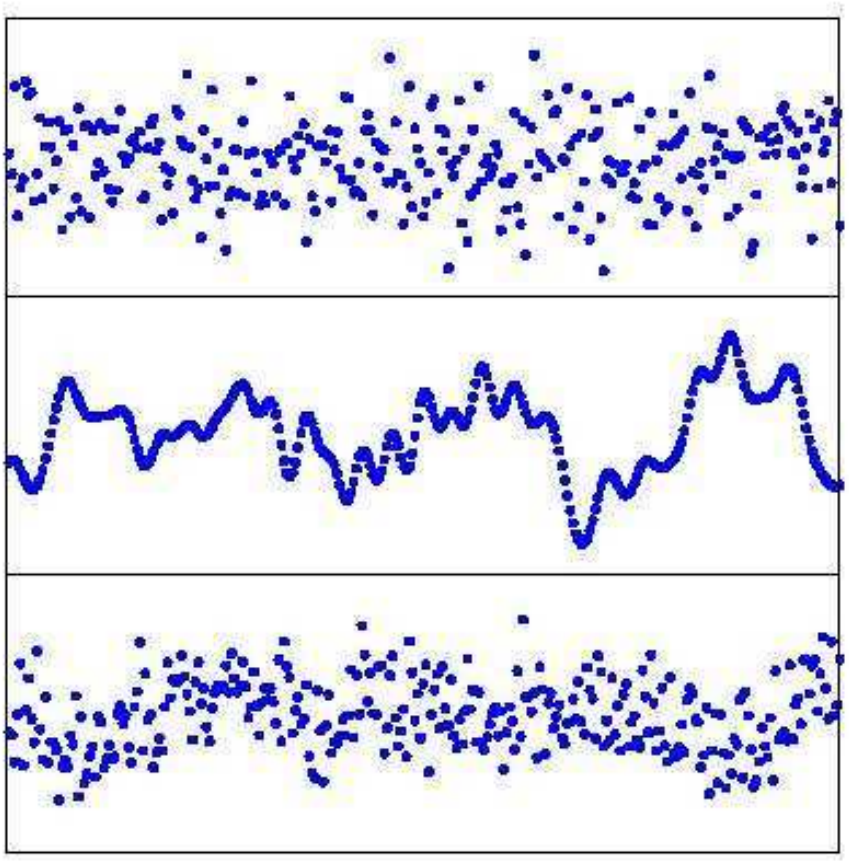


Figure 1.3: **Noise in photometric lightcurves** Figure from Pont et al. (2006) showing lightcurves as they are affected by noise. White noise only (top), red noise only (middle) and white and red noise (bottom).

Sources of red noise are varied and include:

- airmass variations
- changing atmospheric conditions on the small scale - smaller than the telescope aperture
- tracking errors
- focus changes
- flat field errors

Red noise has a significant effect on the detection rate of transiting extrasolar planets. The detection threshold with only white noise is given by Equation (1.2) (Kovács et al., 2002),

$$\alpha \equiv \frac{\delta}{\sigma}\sqrt{nq} \tag{1.2}$$

where σ is the noise level, δ the transit depth,q the fractional transit length, n the number of points (nq is therefore the number of points in transit). α is used as a measure of effective signal to noise. The significance of the signal in the periodogram resulting from running the search algorithm is dependent on this number. When $\alpha < 5$ the signal is noise dependent, while $\alpha \approx 10$ is commonly adopted for transit candidate selection. The detection efficiency was less than expected so that only one quarter of the predicted number of candidates were detected.
Pont et al. (2005) discusses this problem and offers two explanations:

- the ratio of planet and star sizes is smaller than assumed
- the detection threshold is actually higher than previously assumed

In fact red noise affects the detection threshold, uncorrelated noise (white noise) will increase the test statistic thereby decreasing the effective detection threshold when this is taken into account. The real detection threshold, in the presence of red noise,

may well be twice as high. The general effect of this is to increase the baseline of observations required to maximise planet yield. Brown (2003) and Mallén-Ornelas et al. (2003) indicate a baseline of between 20 to 40 days, whereas the actual baseline required is more like 80 to 120 days (Smith et al., 2006). Also wide-field shallow surveys employing aperture photometry are likely to be more susceptible to red noise than narrow deep surveys using difference imaging.

Longer observation periods are not often desired or indeed possible and so other methods for increasing yield were investigated. This led to the development of algorithms to remove red noise from data. There are two methods in use, detailed in §3.8.2 and briefly described here:

- the Trend Filtering Algorithm (TFA) developed by Kovács et al. (2005): a set of template stars is used to calculate coefficients which are combined linearly to reconstruct the lightcurves; systematics that appear in the target lightcurve and the template stars are removed and signals solely belonging to the target remain, and

- the Systematic Removal algorithm (SysRem) developed by Tamuz et al. (2005): lightcurve residuals are solved for linear trends, these are then removed and the algorithm repeats for a specified number of iterations.

Evolving observation strategies, that compensate for systematics and the use of algorithms in the post-reduction stage transit surveys, are now producing many more discoveries of planets. By the end of 2005 only ten transiting planets had been detected, the majority by the OGLE transit survey. Following the recognition of the effect of red noise, seven discoveries were made in 2006. The following year, 18 detections were made and in 2008 there were 19. Table 1.2 gives details of successful wide-field transit surveys. Transits have also been detected by other groups as discussed in §1.4.7.2.

Survey	Telescopes	Size (cm)	Field ($\square^\circ$)	Resolution ("pixel^{-1})	Magnitude Limit	Planets
CoRoT[a]	1	27	3.9	2.5	V < 15.5	5
HATnet[b⋆]	6	11	67	14	I < 12	9
SuperWASP[c]	8	20	482	13.7	V < 11.5	15
TrES[d⋆]	3	10	36	11	V < 12	4
XO[e]	2	20	49	25	V < 12	5
SWEEPS[f]	1	240	0.05	0.05	V < 30	2

Table 1.2: **Transiting planet detections:** Tabulation of planet detections made by transit surveys. For each survey details are given on the telescopes, field of view, magnitude limit and detections made. [a]Bordé et al. (2003) (space mission) [b]Bakos et al. (2004); [c]Pollacco et al. (2006); [d]Alonso et al. (2004); [e]McCullough et al. (2005); [f]Sahu et al. (2006). ⋆ These surveys have telescopes spread longitudinally to increase phase coverage.

1.5 Extrasolar planets and their properties

As the number of detected extrasolar planets climbs (currently more than 300) more information can be determined regarding the statistical properties of the planets. This information is of great use to those working with formation and evolution theories. Selection effects from the detection methods still have a large influence on the statistics obtained, however the parameter space is expanding with more and more different detections each year. Discussed here are properties of the planets and their host stars.

1.5.1 Planet frequency

The increasing number of detected and confirmed extrasolar planets is allowing investigation into the planetary frequency around stars in our galaxy. Data from the CORALIE planet search suggests that 0.8% of non-binary stars have giant planets with semi-major axes < 0.1 AU (Udry and Santos, 2007), and 5.6% of stars have giant planets out to 4 AU. This tallies with the radial velocity results reported by Marcy et al. (2005), wherein hot Jupiters were found around 1.2% of stars from the Lick+AAT+Keck dataset and 6.6% with giant planets within 5 AU. As giant

planets are expected to form preferentially beyond 3 AU this result strongly suggests migration of the planets. The frequency of planets is better determined as a fraction of mass and orbital period with an estimate of detection incompleteness. Monte Carlo simulations (Naef et al., 2005) estimated the fraction of hot Jupiters with periods of less than 5 days as $0.7 \pm 0.5\%$ and the fraction with periods ≤ 3900 days ($\simeq 4.8$ AU) as $7.3 \pm 1.5\%$. Several multi-planet systems have been detected where the planets are orbiting in mean-motion resonance, for example Gleise 876 (Laughlin et al., 2005) and HD128311 (Sándor and Kley, 2006). It is unlikely for planets to actually form in mean-motion resonance so the discovery of such planets gives information that will help constrain migration theories. The number of observed multi-planet systems also gives a lower limit to the true fraction of multi-planet systems since their discovery is a function of selection effects from the detection methods.

Discussions of planetary statistics focus on single stars, however planets have also been detected around multiple-star systems. Predictions by Nelson (2000) indicated that planets should not be found when the stellar separation is less than 40 AU. This has been disproved and planets have been detected around star systems where the separation is as little as 20 AU (Udry and Santos, 2007).

1.5.2 Planet properties

1.5.2.1 Orbital parameters

The orbital parameters of extrasolar planets are interesting in particular due to the differences between detected extrasolar planets and the Solar-System planets. It is of use to investigate these properties in order to determine how planetary systems form. The semi-major axes of extrasolar planets are shown in Figure 1.5.2.1. There is a visible clump at 0.05 AU where the semi-major axis corresponds to hot Jupiters with periods of 3 days. This was completely unexpected. Planets of this size were

expected to form and therefore orbit further from the host star. Butler et al. (2006) indicates that it is expected to be impossible for large planets to form in close, therefore this result indicates migration.

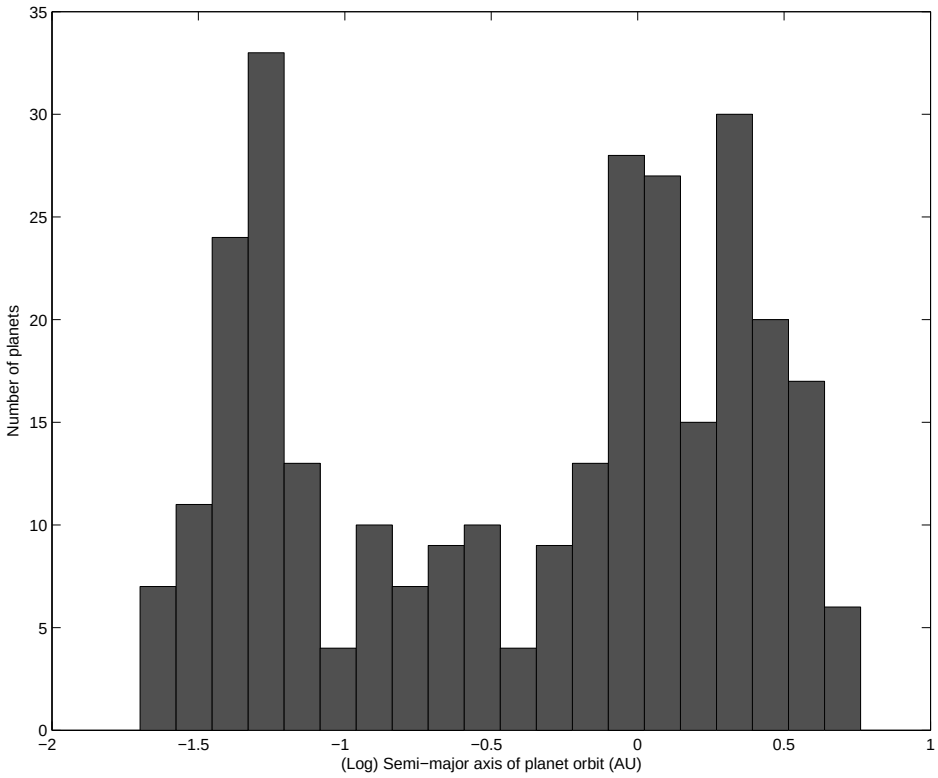


Figure 1.4: **Semi-major axes** Histogram of semi-major axes for planets up to 10 AU, data obtained from the Exoplanet Encyclopedia (Schneider, 2008)

Results from transit surveys give even shorter periods (smaller semi-major axes). This is aided by a selection effect where planets with more transits in the period of observation are more likely to be detected. These planets are known as very hot Jupiters and were originally not detected by radial velocity measurements where the lower limit was 2.5 day periods (Udry et al., 2007). Later detected (O'Donovan et al., 2007), very hot Jupiters were determined to exist independently from the selection effects of transit surveys. Very hot Jupiters appear to be 10 - 20% as

common as hot Jupiters (Gaudi, 2005) but are undetected between 0.02 and 0.03 AU

The natural biases of the most productive search methods have lead to detections of a number of short period planets of all ages. By contrast, direct imaging is sensitive to long period planets, in particular young and massive bodies. This is promising for planetary formation theories. Direct imaging has thus far lead to strong constraints on long period planets and some claimed detections (e.g. Mohanty et al. (2007) and Oppenheimer et al. (2008)). Recently published simulation results by Scharf and Menou (2008) indicate that long-period planets do not oppose the core accretion model and in fact that planet-planet scattering in the early period of the system is probably consistent with the results from direct imaging thus far.

Eccentricities of extrasolar planets cover a large range. Before 1995 all large planets were expected to have circular orbits. This is the case when the semi-major axis is less than 0.1 AU where tidal circularization plays a large part. When the semi-major axis is larger than 0.3 AU, the eccentricities have a flat distribution between 0.0 and 0.8. Non-circular orbits lead to planetary evolution theories that contain perturbative processes such as planet-planet interactions (Chatterjee et al., 2007; Ford and Rasio, 2008; Malmberg and Davies, 2009) and mean-motion resonances which allow planets to end with highly eccentric orbits (Thommes et al., 2008; Raymond et al., 2008). Gas giants in eccentric orbits with small semi-major axes imply the possibilities of induced eccentricities of terrestrial planets in the habitable zone. This may then cause problems with the stability of the system in general, and for habitable planets the stability of liquid water is in doubt with a high forced eccentricity (Adams and Laughlin, 2006).

1.5.2.2 Planet composition and mass

While most of the extrasolar planets detected so far have been gas giants, more recently lighter possibly rocky planets have been detected (Udry et al., 2006, 2007).

Currently there are more than 20 planets around Neptune mass (0.05 to 0.1 M_J). Figure 1.5 shows the rapid fall-off of planets with mass above $3M_J$.

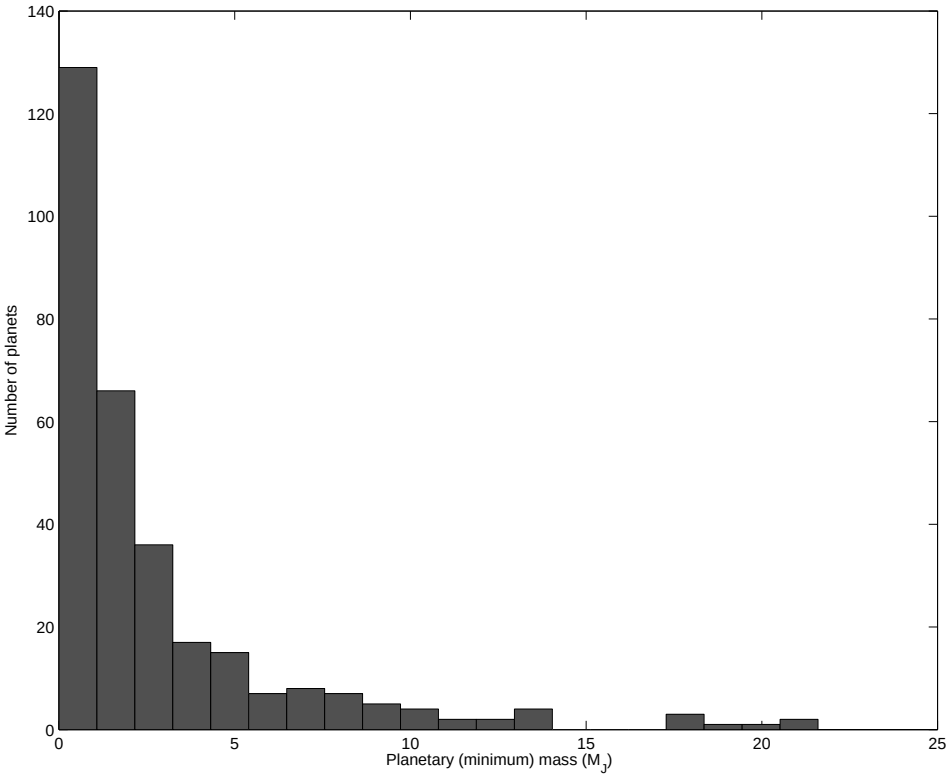


Figure 1.5: **Planet mass** Histogram showing the variation in planet mass, data obtained from the Exoplanet Encyclopedia (Schneider, 2008).

While we discuss here properties of extrasolar planets it is of interest to note that there appears to be a dearth of brown dwarfs with masses between 13 and $80M_J$ with orbits within 3 AU (Grether and Lineweaver, 2006; Patel et al., 2007). This is possibly due to a difference in the formation methods of stars and planets and there may be a true lack of companions in this mass range.

As yet there have been no detections of Earth-mass planets but with the increasing number of planets detected each year and the large numbers of surveys in motion - including those specifically designed to detect Earth-like planets it is not

impossible to believe that the detection of an Earth-mass planet will happen within the next few years.

With the addition to the database of a large number of transiting planets confirmed by radial velocities, mean densities and therefore compositions of extrasolar planets are beginning to be investigated. The theoretical predictions have already come under fire with several planets detected with unexpected densities, for example:

- HD209458b (Charbonneau et al., 2000) with density less than that of Saturn which has the lowest density of the giant planets in the Solar-System,
- TrES-4 (Mandushev et al., 2007) which has a large radius and extremely low density,
- HD149026b (Sato et al., 2005) which was modelled to have a large dense core and
- Gliese 436b (Gillon et al., 2007) suspected to have a main constituent of water ice as well as an outer layer of H/He in order to account for the small radius.

These and other planets with unusual densities have had large effects on the theoretical models. The planets which do not fit with known models will help constrain evolution and migration theories. For example Gliese 436b borders on the change from a gaseous to icy planet but is too close to its host star for the planet to have formed there. It must instead have formed out past the "snow line" and migrated inward.

1.5.3 Stellar properties

1.5.3.1 Stellar mass

In studying the stellar mass of extrasolar planet host stars, Johnson et al. (2007) determined that there is an observed lack of close-in planets around intermediate-mass stars and that $> 7\%$ of stars have giant planets at greater distances. There is a slight bias towards hot Jupiters being more common around heavier stars, less than 2% are found around 0.1 - 0.6 $M_{\odot}$ M dwarfs (Bonfils et al., 2007). Lower mass stars have lower mass planets on average (Bonfils et al., 2007) with 30% of planets with $m\sin i \leq 30 M_{\odot}$ occuring around M dwarfs. This is explained by core accretion (Ida and Lin, 2005; Laughlin et al., 2004), which predicts inhibition of the formation of giant planets around low-mass stars.

Burkert and Ida (2007) determined that there is evidence for a gap in semi-major axis distribution between 0.08 AU and 0.6 AU for stars with mass greater than $1.2 M_{\odot}$. This gap was reproduced in simulations and has been ascribed to the shorter length of time involved in the depletion of disks around intermediate-mass stars.

1.5.3.2 Stellar metallicity

Investigations of stellar metallicity show that more metal-rich stars are more likely to host Jupiter mass planets, Figure 1.6. Studies by Fischer and Valenti (2005) show that more than 25% of hot Jupiters have been found when the host star has metallicity of [Fe/H] > 0.3. This bias is not seen in lower-mass planets (Udry et al., 2006). There is a trend for high-core mass around metal-rich stars (Burrows et al., 2007).

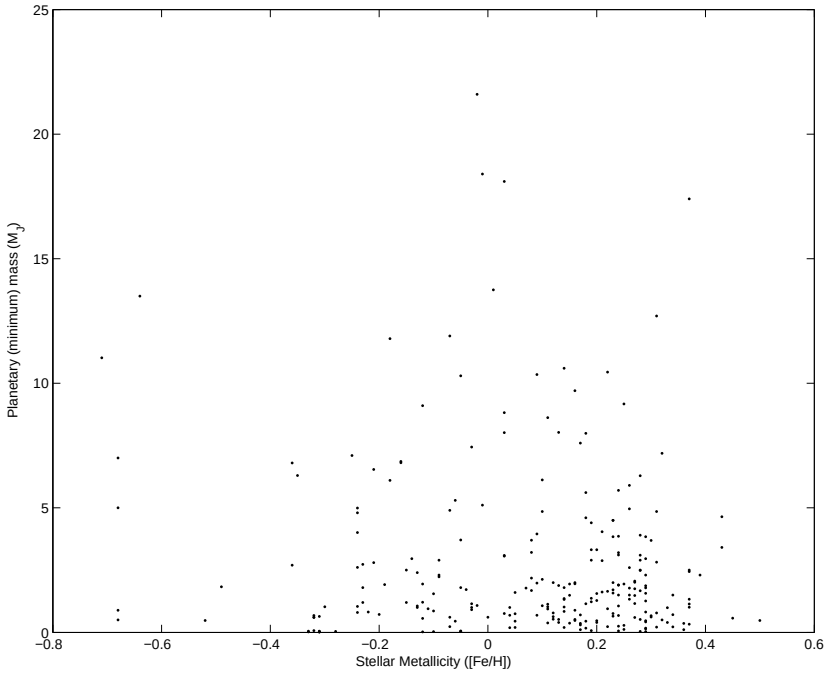


Figure 1.6: **Host star metallicity** Planet mass as it varies with stellar metallicity, data obtained from the Exoplanet Encyclopedia (Schneider, 2008).

1.6 History of observations of variable stars

A variable star is, quite simply, a star whose observed brightness changes with time. There are many causes for variability ranging from starspots (as seen in the 11 year sunspot cycle) to supernovae (catastrophic explosion of the star). A large number of the variations are periodic and are caused either by internal changes in the star (such as pulsations) or by external factors (such as a companion eclipsing the primary).

Variable stars have been observed for hundreds of years. The first periodic variable was detected (and recorded) in the 1600's. This star was a pulsating star where the variations are caused by internal changes in the star. Variables of this type are known as intrinsic variables. The second periodic variable star was discovered in 1669 but the correct reason for its variability was not determined until 1784 when John Goodricke determined that the variations were due to the periodic occlusions of binary partners. The variability in this case was caused by external factors (extrinsic variability). By 1786 a total of 12 variable stars were known, and this number has increased rapidly since especially with the advent of photography. Photography allowed direct comparisons of the brightness changes in a star. The Galactic Catalogue of Variable stars now includes more than 50,000 variables with another 10,000 pending confirmation.

1.7 Variable star observation methods

Studies of variable stars involve observing and taking measurements of the star and then plotting the changes with time. Regular variables generally have a well defined period and amplitude though many variable stars have periods and amplitudes which will themselves vary with time.

The two primary methods for variable star study are spectroscopy and photometry. Photometry is the traditional method of observing brightness changes in the star, while spectroscopy examines changes in the spectrum of the star.

Spectroscopic observations split the light from a star into a spectrum, the spectrum can then be examined and compared to previously known measurements. From the spectrum of star we can obtain data such as the temperature and luminosity class, whether the star is a single or binary star, radial velocity variations, temperature changes, and the unusual variations that occur in a single part of the spectrum. A common use of spectroscopic measures is determining the radial velocity of a star. The changes in wavelength of particular parts of the spectrum allow a measure of the speed the star is travelling. Spectroscopic binary stars are identified by the changes in the star's radial velocity as the binary pair rotates around a common centre of mass. Problems with spectroscopy are that in most cases only a single star can be examined at one time. A large number of stars are also too faint for current spectroscopic measurements. Spectroscopic measurements are often used in followup of previously detected variable stars to obtain more information.

Photometric observations take measures of the brightness of a star. The brightness is plotted against time and the resulting lightcurve examined for periodic changes. Identification of types of periodic variations is performed by examination of the lightcurve period and morphology. Less information is obtained from photometric measurements but if a set of different filters is applied to the observations more information about the star type can be determined. Photometric observations can be performed either on a single star or on a large number of stars. The advent of digital technology and CCD cameras had made photometric observations simple. Fainter stars can be detected by photometric instruments that spectroscopic and thus variable stars can be detected in distant places.

1.8 Motivations for variable star research

Studies of variable stars improve our knowledge of the universe. Variable stars are useful for learning distances, about the way stars evolve, about star formation and

destruction. Binary stars are of particular interest in the way they evolve with time and effects of a faint or low-mass companion on the primary allow measurements to be taken of stars that are often undetectable.

1.9 Variable star types

Variable stars fall into two types: intrinsic and extrinsic variables. The extrinsic class contains any variable whereby the variation is caused by external factors. This includes variable stars such as eclipsing binaries where the variation is caused by binary companions passing rotating around each other in the line of sight; or rotating variables where the variations are related to the stars rotation, either by speed of rotation causing distortions or by changes in the stars surface with longitude. Intrinsic variables have variations caused by internal changes in the star. The most common is pulsations of the star which can be found in many forms. Stars whose surfaces erupt with flares or mass ejections are intrinsic variables. Explosive or cataclysmic variables such as supernovae and novae are also classed as intrinsic variables. These types are most often non-periodic.

There are a large number of different types of variable stars. They range from single event supernovae to periodic variations on the scale of decades. This research is most likely to detect short period variations in particular those caused by eclipsing binaries. These have the closest characteristics to transits and so a project suited to transit detection will also result in a large number of eclipsing binaries. BY Draconis stars may also be detected but are commonly confused with contact binary lightcurves so identification of these is less likely. Other possibilities are RR Lyrae stars which have periods of less than a day with amplitudes between 0.5 and 1.5 magnitudes. Similarly δ Scuti stars should be detectable. It is also possible that short period Cepheids may be detected. Here are discussed these types in more detail.

1.9.1 Eclipsing binary stars

Eclipsing binary stars have the orbital plane in the observers line of sight. This causes periodic variation in the light signal from the system due to the changes as the components pass in front and behind each other. There are three main types of eclipsing binary stars and they are characterized from the shape of their lightcurves.

1.9.1.1 Algol type eclipsing binaries

Algol types (EA) named after Algol (β Persei) are characterized by the light remaining constant between the eclipses. The amplitude of the eclipses ranges widely and the primary and secondary eclipses may be comparable in depth or extremely unequal, in rare cases the secondary eclipse may be so shallow as to be undetectable. EA binaries are produced when both components of the system are approximately spherical - one component may be non-spheroidal only if it contributes little light to the system. Orbital periods of EAs range from a few hours up to several decades and the sharp eclipses allow accurate measures of the periods. The stellar radius/radii of EA binaries are a relatively small fraction of the separation of the stellar components.

1.9.1.2 β Lyrae type eclipsing binaries

β Lyrae type (EB) binaries have a continuously varying lightcurve between the eclipses. The primary and secondary eclipses vary significantly in depth and the orbital periods are always longer than one day. One or both of the components will be highly elliptical and fill its Roche lobe. The stellar components have spectral types B or A. The lightcurves of EB binaries are far more sinusoidal than the EA type. Sometimes the lightcurve of a highly distorted star may appear to be that of an EB binary.

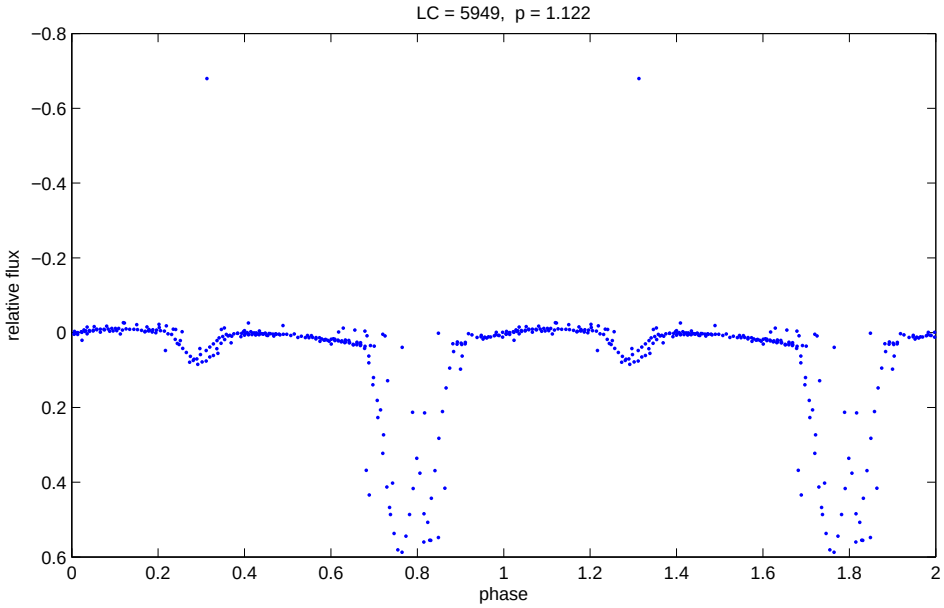


Figure 1.7: Sample lightcurve from the dataset containing an Algol type eclipsing binary with period of 1.122 days.

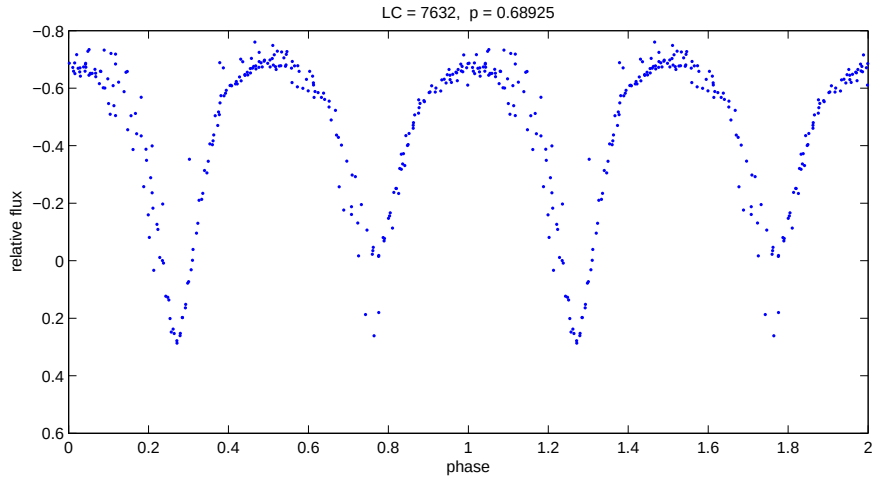


Figure 1.8: Sample lightcurve from the dataset containing a β Lyrae type eclipsing binary with period of 0.68925 days.

1.9.1.3 WUma type eclipsing binaries

WUma type (EW) binaries are characterized by continuous variation of the lightcurve due to both the eclipses and the changing aspects of tidally distorted stars. The minima of the lightcurves are of almost the same depth and the assumption is that both stars are in contact. The minima would indicate that the stars are of similar luminosity but this can also be explained by the massive component transferring luminosity to the less massive by means of a common envelope, and thus equalising the surface temperatures. The EW systems are of low-luminosity with periods of less than one day. The period changes with time as a result of the the ongoing mass changes between the two components. EW type binaries are also designated contact or close eclipsing binaries.

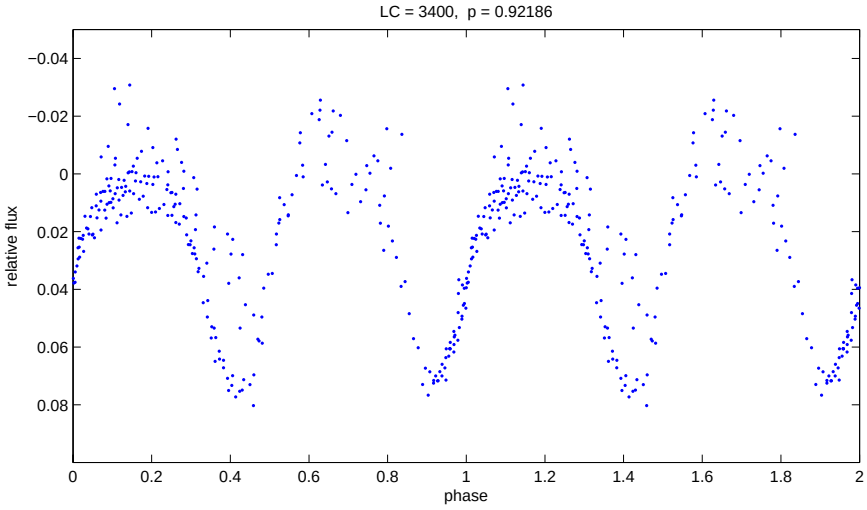


Figure 1.9: Sample lightcurve from the dataset containing a WUma type eclipsing binary with period of 0.92168 days.

1.9.2 Cepheid type stars

Also known as δ Cepheids, Type I Cepheids or classical Cepheids, Cepheids are strictly periodic with periods for the most part between 1 and 50 days. The

lightcurve varies smoothly from shorter to longer period stars, with increasing width of the maxima. At short periods, bumps may be seen on the falling branch of the lightcurve, while at longer periods the bumps occur on the rising branch. At longest periods the lightcurve becomes more sinusoidal. There is a strict relation between the mass of the Cepheid and the log of the period. As Cepheids are young, massive stars they are bright and can be detected at large distances. Cepheids are used as distance indicators as the mass and luminosity can be determined from the period. Cepheids pulsate mostly in the fundamental mode though a few cases of lower amplitude Cepheids are known to pulsate in overtones with periods from 1.5 to 4 days.

1.9.2.1 Type II Cepheids

Type II Cepheids are commonly considered low-mass analogues to classical Cepheids and are not well understood. Masses of Type II Cepheids are around $0.6M_{\odot}$ with periods ranging from 0.75 to 40 days. There is a general dependence on period but due to variation in periods and lightcurves from cycle to cycle this is not well constrained. Like Cepheids they are radial pulsators with the variation caused by shock wave propagation. Some theories consider that Type II Cepheids may be formed by the merging of two components of a binary and that is is why they are are so hard to categorize.

1.9.3 RR Lyrae stars

RR Lyrae stars are commonly called cluster variables due to their prevalence in globular clusters. They are members of the halo and old disk populations. Radial pulsators with periods of 0.2 to 1 days, they are most likely a different evolutionary stage of the same star type as Type II Cepheids and like Type II Cepheids spectra of RR Lyrae stars show evidence of shock waves propagating through their atmospheres. RR Lyrae stars span a wide range of metallicities from $[Fe/H]\odot$ to

hundredths of [Fe/H]$\odot$, and there is a dependence of absolute magnitude on metallicities. RR Lyrae stars can be used as distance indicators and are particularly useful for this as they are old stars and may be found in systems where there are not the younger Cepheids. They are found on the horizontal branch of the H-R diagram.

Their are two main groups of RR Lyrae stars:

- RRab type: high amplitude variation (1 mag is common), asymmetrical lightcurves (steep rising slope), pulsate in the fundamental mode and have longer periods (0.4 - 1.0 days),
- RRc type: lower amplitudes (0.5 mag), more sinusoidal lightcurves, pulsations in the first overtone and shorter periods (0.2 - 0.5 days).

Double-mode RR Lyrae stars (now commonly known as RRd) show variations in shape and amplitudes of lightcurves between cycles. There is also a long-term variation of lightcurves in the range of 20 to 200 days known as the Blazhko effect, and in some cases this modulation is further modulated on even longer time scales. It is plausible that the Blazhko effect could be detected in RR Lyrae stars found in this data set.

1.9.4 BY Draconis type variables

BY Draconis type stars are rotating variables where the variations is produced by axial rotation of non-uniform stars. The variation is due to chromospheric activity (star spots etc) but the variable may be a single or binary system. In the case of a binary system the variability occurs between eclipses in the form of a star spot wave. BY Draconis variables are late dwarf type stars.

1.9.5 Cataclysmic variables

Cataclysmic variables are characterized by long periods of quiescence with dramatic outbursts of large amplitude. Supernovae and novae are both classed as cataclysmic

variables as are various other types. It is unlikely that a detection of a cataclysmic variable in outburst will be missed due to the large change in amplitude but they can also be detected in quiescence with low amplitude spikes in their lightcurves.

1.9.6 δ Scuti stars

δ Scuti stars are type A or F stars with small amplitude variations. The amplitudes range from 0.8 mag down to several thousandths but most are found with amplitude around 0.2 mag. δ Scuti stars occupy the instability strip and are radial pulsators. Periods of δ Scuti stars are less than 0.3 days. They have multiple radial modes which causes multi-periods variations and non-sinusoidal or skewed sinusoidal lightcurves.

1.10 This work

This work was originally started in order to take advantage of observations obtained by the PLANET collaboration on the 2.2m telescope at La Silla, Chile in 2002. The initial aim was to search the lightcurves from this dataset for transits. The high-precision photometric data of a wide field was suitable for a transit search and for detection of variable stars. The observation field is crowded with a large number of stars and reaches to a magnitude of R $\sim$ 24.5. The aims of this research are:

- to reduce the dataset to high precision lightcurves
  - development of a image reduction system for the images
- to search the dataset for transit candidates
  - implement code to run transit detection algorithms
  - simulations of transiting planets to test the algorithms
  - determine possible candidates
  - characterize the candidates

- to search the dataset for variable stars
  - implement variable star detection algorithms
  - simulations of variable stars to test the algorithms
  - determine possible candidates
  - characterize and build a catalogue of variable stars in the field
  - determine the binarity of the field
  - identify variable stars with unusual properties

Chapter 2 summarizes the observations used in this research and the expectations from the dataset. Chapter 3 discusses the reduction process used for the dataset including precision of the results. The resultant lightcurves and systematics in the lightcurves are also discussed and dealt with. Chapter 4 discusses the possible search algorithms in brief and those search algorithms used in both the transit search and variable star search are discussed in detail. Testing the algorithms is discussed and thresholds for candidature is determined. The initial results from the search algorithms are also discussed here with the transit candidates detailed in Chapter 5. The variable stars discovered in this research are described and catalogued in Chapter 6. The conclusions and discussions of further work on this dataset are contained in Chapter 7.

Chapter 2

Observations

The target field for this research is close to the galactic plane (Galactic latitude $-2.2779°$) in the Norma Spiral Arm. Figure 2.1 shows the approximate location. The coordinates of the centre of the field are: right ascension $16^h19^m20^s$, declination $-53°26'40''$ or galactic longitude $330.9368°$, galactic latitude $-2.2779°$. The field is relatively crowded and we expect around 300,000 stars with usable photometry. There are few bright stars in the Norma constellation and the field contains only a few known variable stars (see §6.3.2). There is also a planetary nebula, a molecular cloud, several X-ray sources and a large number of IR sources in the field (Figure 2.2).

The microlensing followup group Planetary Lensing and Anomalies Network (PLANET) observes at the La Silla observatory on a regular basis on the 1.5 metre Danish telescope. In 2002 the PLANET collaboration instead had time on the larger 2.2 metre telescope nearby and using the wide field camera on this telescope obtained additional images in the hopes of detecting extrasolar planetary transits.

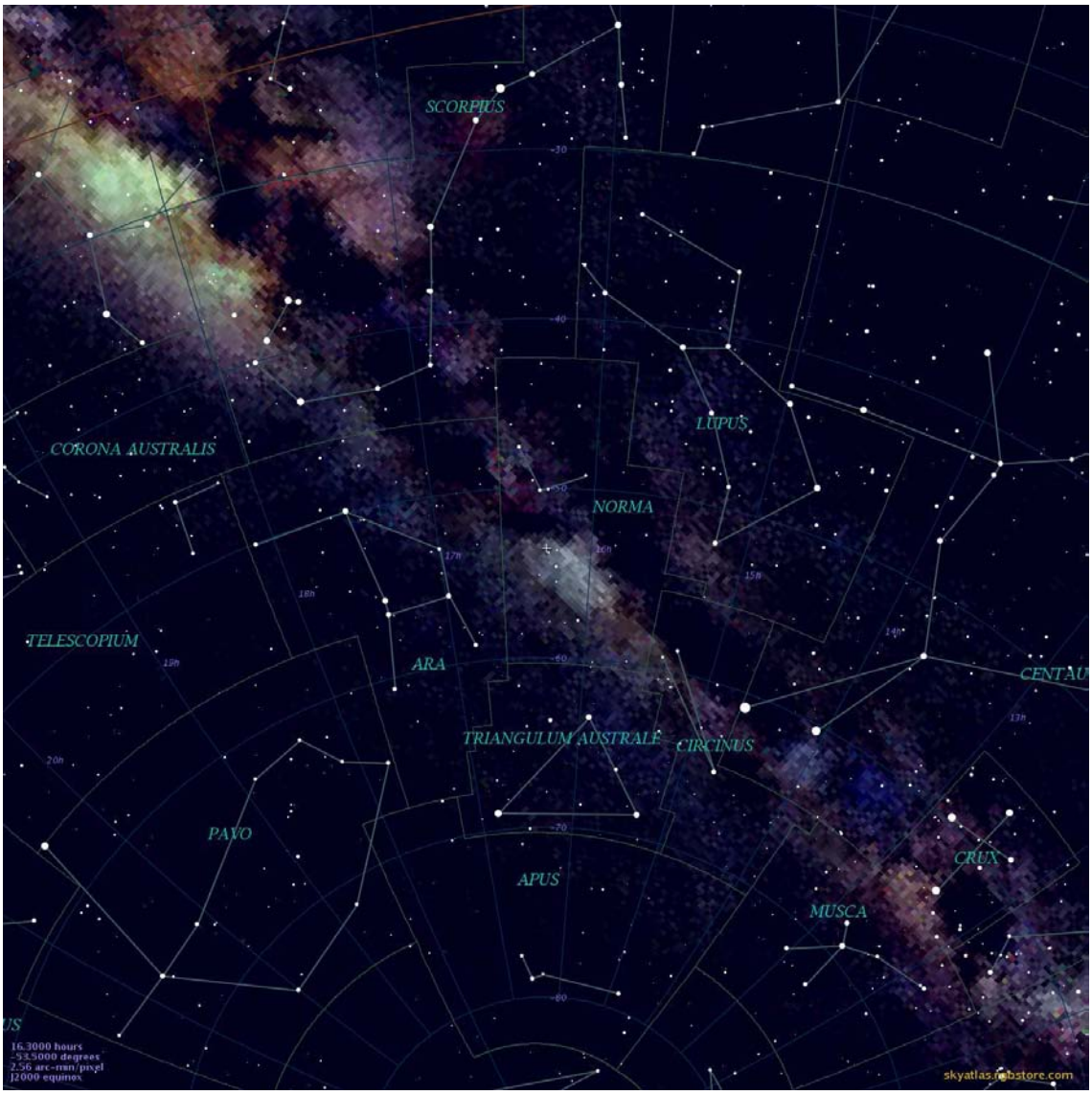


Figure 2.1: **Observation field in the Milky Way** Image of the Milky Way and constellations. The observations used in this research are acquired from approximately the marked + in the constellation Norma.

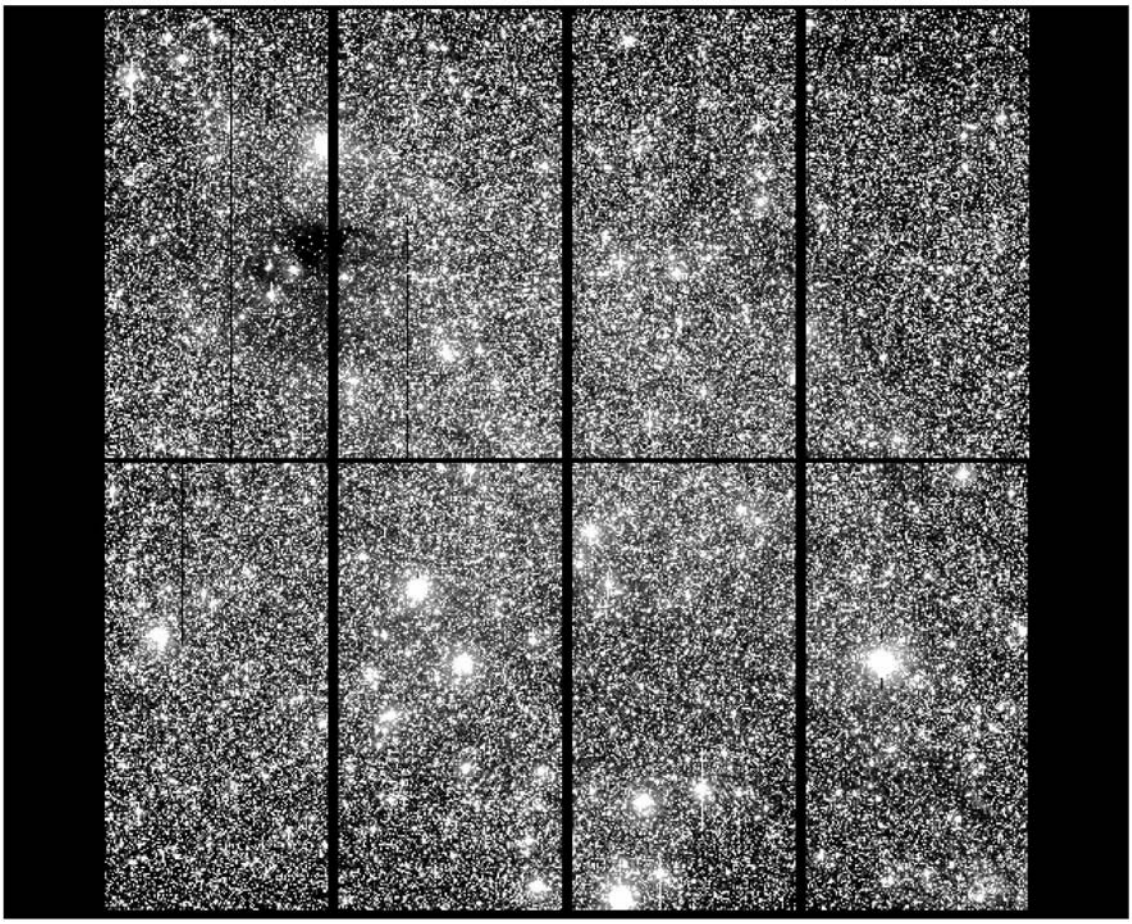


Figure 2.2: **Mosaic CCD image** Mosaic CCD image from the WFI camera showing the entire data field (R band).

2.1 Telescope

The telescope used for observations for this research is the 2.2 metre telescope at the European Southern Observatory (ESO) site at La Silla in Chile. The telescope is officially owned by the Max-Planck Institute but based on the ESO site. The observing time is shared by ESO and Max-Planck observers. The telescope, in operation since 1984, is on an equatorial mount and has an f/8 focal ratio at Cassegrain focus. The primary 2.2 metre mirror is a hyperbolic made of Zerodur, as is the 84.4 centimetre secondary. There are three instruments used on the telescope. The Wide Field Imager (WFI) (§2.1.1) which was used for this research; a fibre-fed, bench mounted, environmentally controlled, high resolution Èchelle spectrograph, FEROS; and a Gamma-Ray Burst Optical/Near Infrared detector called GROND designed to observe the afterglow of Gamma-Ray Bursts and other similar events in multiple filters simultaneously.

2.1.1 Camera

The WFI (Baade et al., 1999) camera is permanently mounted at the Cassegrain focus of the 2.2 metre telescope, with a $34'$ by $33'$ field of view. The camera is constructed from a four by two mosaic of 2k by 4k CCDs giving a $0.238''$/pixel scale. The camera has excellent sensitivity from the atmospheric cutoff into the near infrared (approximately 350 nm to 1 μn). There are 40 filters available for the camera and those that were used in this research are discussed in §2.1.2. Fig. 2.2 shows a mosaic image of the observation field taken by the WFI camera in the R band.

2.1.2 Filters

The filter used for the majority of the science observations is a filter in the R band. The central wavelength is 651.725 nm with a full width half maximum of 162.184

nm and peak wavelength of 668.5 nm. The percentage transmission at the peak wavelength is 93.9. The R filter used in these observations is a close match to the standard Cousins R filter. A few images were also taken in the B and I filters, with peak wavelengths of 475.0nm and 760.0nm respectively. These images will be used to construct colour-magnitude diagrams of the field in order to make appropriate identification of variable stars and transit candidates (see §3.9.3).

2.2 Pilot dataset

In the winter observing season of 2002 the PLANET collaboration was granted time on the 2.2m telescope at the ESO observatory in La Silla, Chile. PLANET performs followup observations on microlensing objects detected by the microlensing survey teams and in general has made use of the 1.5m telescope at La Silla. When time was awarded on the 2.2m telescope the attempt was made to obtain observations for use in transit detection to take advantage of the larger aperture and field.

The observations were compiled over much the same timescale as normal microlensing observations. Microlensing observations are collected on a one to two month timescale with widely spaced observations for the majority of the time and a few days of closely spaced observations over the peak of the microlensing event (see §1.4.2 for more information). Three fields were observed in 2002 and each field has between 50 and 70 images spread over around 6 weeks. The initial reduction resulted in several hundred thousand lightcurves. After initial attempts to search for transits it was found that the current data was not particularly suitable. A large portion of each dataset is confined to one or two days in the centre of the time frame and this makes the observations less than ideal for transit detection unless a transit happens fortuitously to occur within that time. Attempts were made to overcome this problem and these are discussed following.

2.2.1 Temporal binning

Temporal binning is the method of taking a weighted average of data points within a short time scale. It results in a smaller number of data points but improves the standard deviation of the dataset. As a planetary transit generally lasts for several hours, analysing the data after taking a weighted average of the points within each hour is a valid procedure. Possible transit candidates stand out from the other stars more easily in binned data than in un-binned data. Figure 2.3 shows the change in standard deviation pre- and post-binning. The solid line shown on the two graphs is the theoretical transit depth for a hot Jupiter at different magnitudes. This shows a sample of the sensitivity of the data to detection of a planet orbiting stars of different magnitudes. The larger number of data points falling below this line the larger chance of detection of a transit. Binning the data into hourly blocks makes for a large decrease in the standard deviation as can be seen in Figure 2.3; overall the entire group of stars moves downward on the plot. The number of stars that fall under the theoretical detection line increases dramatically, clearly this is a sensible avenue to follow when there are at least eight data points per hour.

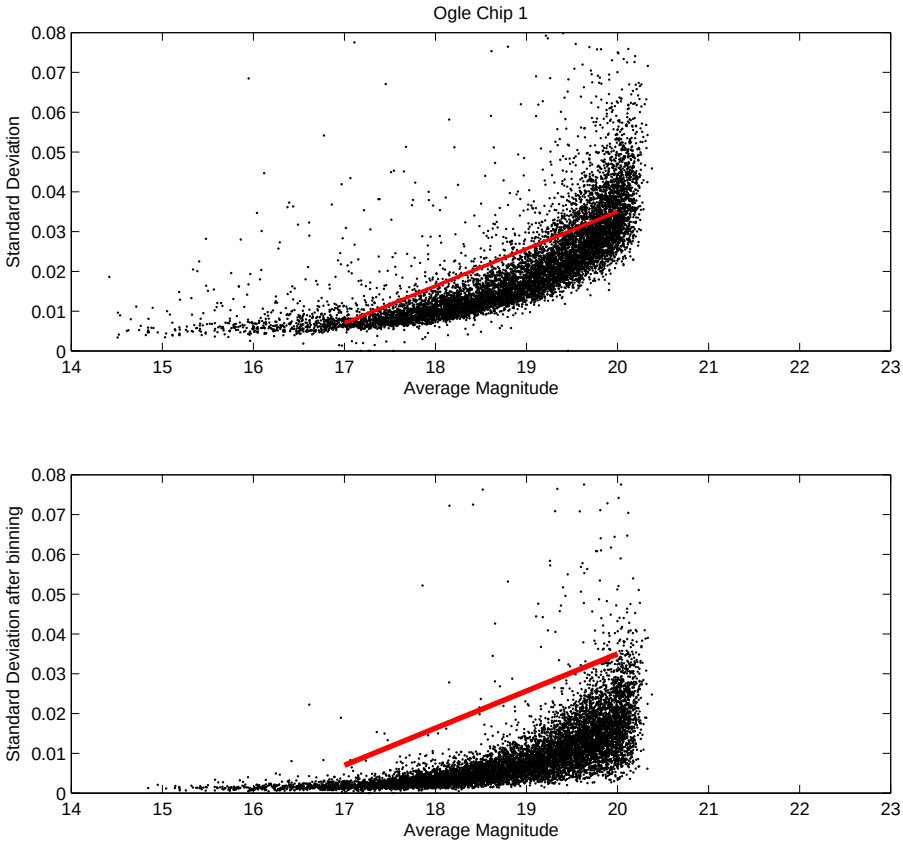


Figure 2.3: **Theoretical extrasolar planet detection** The change in standard deviation of the pilot data as it compares to detection levels pre and post-binning of the datapoints into hourly boxes. The number of stars falling under the theoretical detection line (bold line) increases substantially with hourly binning.

2.2.2 Precision of data

From the pilot data of one chip of one field that contains 80,926 detected stars, there are 28,066 with photometric precision (binned data) better than 0.5%, (Table 2.1). Most of the stars fall into the uncalibrated magnitude range of 19 to 20. The frequency of hot Jupiters in the solar neighbourhood is 0.75% (Butler et al., 2001), with a geometric transit probability of 10%. The likelihood of an observable transiting planet around a given star is therefore 0.075%.

σ	15-16 mag	16-17 mag	17-18 mag	18-19 mag	19-20 mag	Total
0.5%	387	2019	6528	8361	10771	28066
0.6%	387	2026	6697	10165	14178	33453
0.7%	388	2027	6764	11371	17278	37828
0.8%	389	2030	6805	12125	20025	41374
0.9%	389	2032	6834	12630	22467	44352
1.0%	389	2032	6846	12959	24604	46830

Table 2.1: **Precision of data:** Tabulation of the number of stars with standard deviations less than 0.5% to 1% in magnitude blocks.

With the pilot data we have high sampling over two days of observations and using a Monte Carlo simulation we find the fraction of transiting planets with one observable transit in this dataset to be 0.1. This gives us a detection probability of 0.0075% per star and from the 28066 stars with the best photometric precision we would expect to observe 2.1 transiting planets. If we were to do a 20 day campaign, (§2.3), then the fraction of stars with three transits - required for detection, is 0.17 which leads to observing 3.5 transiting planets. However if we require only two transits the fraction of detected transits increases to 0.5 and we would expect to find 10.5 planetary transits out of 28066 stars. If we aim to detect only two transits, the candidate stars will require later followup observations to make an accurate determination of planetary existence. Given that the results from the original set of data collected in 2002 were not as suitable for transit searches as was hoped, it was proposed that an additional amount of telescope time was acquired with better

temporal spacing. A proposal was made in association with Cristina Afonso of the Max-Planck Institut fuer Astronomie to use some of their dedicated time on the ESO 2.2 metre telescope at La Silla. The rationale for the proposal was as follows.

2.3 Observations

For a dataset to be suitable for the purposes of transit detection there are several criteria that need to be fulfilled (Gillon et al., 2005; Deeg, 1998). To detect a planetary transit of the type described in §1.4.7 one requires a time span long enough to include 3 to 4 transits. In order to search for planets with a period of 3 to 4 days, this requires approximately 16 days of continuous observing. A single transit has a duration of a few hours - to detect this in a lightcurve one requires a continuous well-sampled lightcurve. There needs to be a number of points within the transit well so that the transit is detected. The required precision for the observations was a large factor in determining the number of observations desired. In order for the transit depth resulting from a hot Jupiter to be detectable within the lightcurve, the precision of the data points must be 0.7% for an F0, magnitude R = 17 star, and 4.4% for an M0, magnitude R = 21 star, (see Figure 2.4).

The fraction of dwarf stars with hot Jupiters is assumed to be 1%. The probability that the transiting star is geometrically aligned with the observer is approximately 10%. Monte-Carlo simulations gave the fraction of stars with three or more detectable transits in the observation period as 0.03 for 16 days. This leads to a total transit probability of 0.3 transits per 10,000 stars. Given that we expect approximately 300,000 stars with the appropriate photometry this leads us to expect approximately nine detectable planets (see §1.4.7.3 for more detail). It is however unlikely that all such planets will be detected. Despite good photometry the stars may well be too noisy or variable to enable a smooth detection. Our transit search

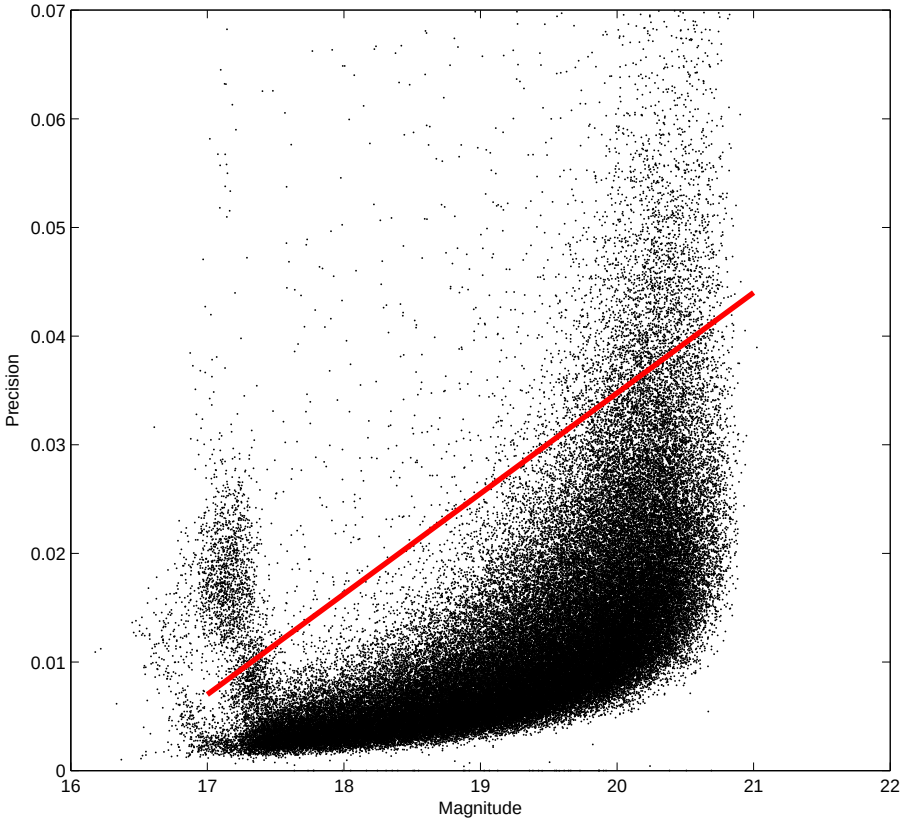


Figure 2.4: **Required data precision** for detection of planetary transit: the standard deviation of the data points plotted against the average magnitude for each star. The line shows the theoretical detection limit of a hot Jupiter - 0.7% for an F0, magnitude R = 17 star and 4.4% for an M0, magnitude R = 21 star. (Note that the bright magnitude low precision tail is caused by stars at the saturation limit.)

proposal was designed to search for hot Jupiters, with closely spaced observations over approximately 16 days of one of our original fields.

In response to our proposal, 102.5 hours of observing time were granted. Observations were to be taken in service mode using the R filter and approximately an 8-9 minute integration to achieve a signal to noise of 100 on a K-type star down to R = 21 mag. Approximately 624 measurements were expected, giving 174,000 stars with the appropriate photometric precision.

The observations were acquired in July - August 2005 and resulted in 204 R filter images taken over a one month period with variable seeing plus three exposures in each of the B and I filters. The acquired observations were fewer than expected but are still suitable for analysis albeit with a reduced detection sensitivity. The expectation, given approximate transit duration as compared to an average expected period, is that a transit will contain around ten points. This is a reasonably significant number of data points and is promising for detection probabilities.

2.4 Dataset

The total dataset (including the pilot data from 2002) is comprised of 267 images, though a few were bad enough to be removed from the dataset prior to processing. For the most part the stellar images have a full width half maximum of between 5 and 8 pixels (see Figure 2.5). The sky background in the useful observations ranges from 1300 counts to approximately 4000 counts. The median sky background count is around 1800 counts (see Figure 2.6).

The observations are tabulated with filter and date in Table 2.2. Details of the reduction are contained within Chapter 3.

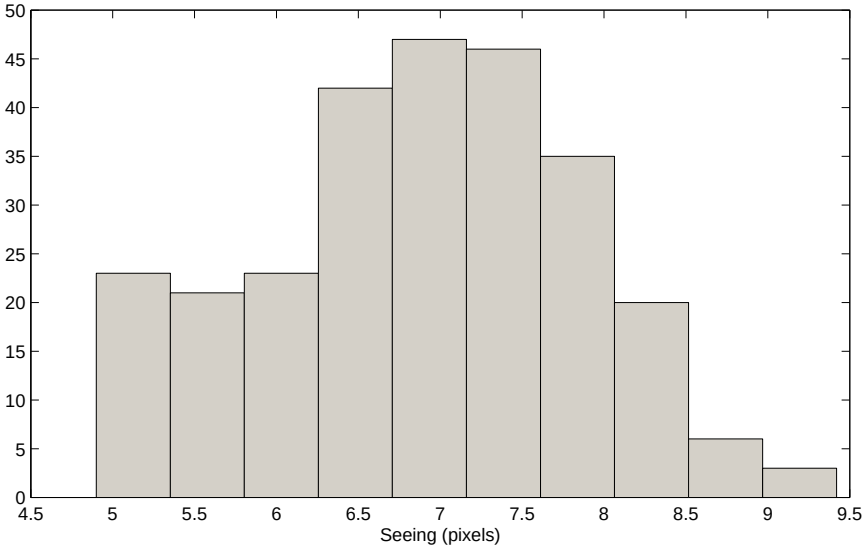


Figure 2.5: **Seeing:** Histogram of seeing in the observations.

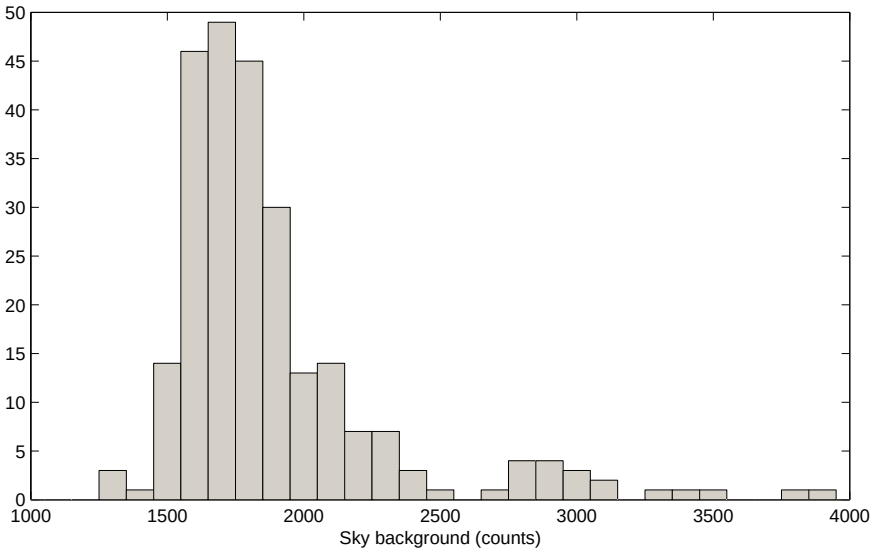


Figure 2.6: **Sky background:** Histogram of sky background values in the observations.

Year	Date	Filter	Number	Exposure time
2002	25 June - 15 August	R	63	600
2005	26 July - 25 August	R	204	720
2005	26 July - 25 August	B	3	720
2005	26 July - 25 August	I	3	720

Table 2.2: **Total observations:** Tabulation of the number of observations collected in 2002 and 2005 used in this research

Chapter 3

Reduction

Astronomical image reduction, in our context, is the process of converting raw CCD images to lightcurves for all the stars of interest. This involves several preprocessing steps, e.g. bias removal and flat-field division followed by the extraction of photometric measurements. There are three possible methods that can be employed to do this: aperture photometry, point spread function (PSF) photometry or image subtraction (also known as difference image analysis).

Aperture photometry involves counting the number of photons collected in a circle around the star. An annulus around this circle is used to make an estimate of the sky background included in this count and this is then subtracted off. Aperture photometry works well for isolated stars but is not useful for crowded fields with large numbers of overlapped or conjoined stars. Aperture photometry is a fast method as the major calculation only contains a single sum of points.

PSF photometry, also known as profile fitting, involves modelling the light from a star usually by a Gaussian point spread function. The PSF represents the spreading of the light from a point source star as caused by the atmosphere, telescope, filter and detector. Given an initial position of a star the PSF is centered and then scaled to the actual profile. Each star may overlap with that of another but the PSF is used to estimate what proportion of the light in each pixel originates from

each star and therefore can be used in crowded fields where aperture photometry is not possible.

Image subtraction, also known as difference image analysis, involves subtraction of a reference frame from each image and then calculation of the residual brightness of target stars by means of an aperture measurement or PSF fit. The reference frame is convolved to match the seeing of each individual image to ensure the star sizes are the same. The reference frame is then subtracted off and the star counts calculated. Image subtraction is suitable for crowded fields. The sky is generally estimated across the field but can be estimated in the area around the star. Image subtraction is not a fast method - convolution of the kernel prior to subtraction of each image takes a large amount of time and can be very hard to do accurately with large images. The nature of the observations used here mean that image subtraction is the only suitable method for reduction. The field is very crowded and many of the stars overlap with others.

The data collected in 2002 was originally reduced by Dan Bramich (University of Saint Andrews) with his own code, and was available in the form of lightcurves with the pre-processed images alongside. In order to test the precision of the data and the then newly developed ISIS pipeline (see §3.1.4) designated for this research, the 2002 data was re-reduced when the proposal for a new set of observations in 2005 was submitted. This proved that the new reduction technique was superior to the original method. As a consequence when the 2005 data was available both that and the original unprocessed frames from the 2002 observational period were reduced together using the techniques outlined in this chapter.

3.1 Software

3.1.1 IRAF/PyRAF

The Image Reduction and Analysis Facility (IRAF) (Tody, 1986) is a data reduction and analysis system used by many astronomers to aid in analysis of observational data. PyRAF (De La Peña et al., 2001) is a newer, easier to use and more interactive interface to IRAF based on the PYTHON command language. The development of PyRAF has enabled more flexibility in script writing and graphical interaction. The PyRAF interface allows the combination of IRAF tasks with the more expansive capabilities of Python coding. We use here the PyRAF capabilities to perform the pre-processing on the images. It is also used to determine coordinates and reference magnitudes from the stars in the reference frame (see §3.4.1).

3.1.2 DS9

SAOImage DS9 (Joye and Mandel, 2003) is an image viewer created by the Smithsonian Astrophysical Observatory which provides image support for various astronomical analysis programs including IRAF and PyRAF. In conjunction with IRAF or PyRAF, DS9 can be used to identify stars from a particular coordinate list and in interactive mode allow extra unidentified stars to be added to this list. This is the main way in which DS9 has been used in this research, but we have also used DS9 on its own to test the subtraction and photometry. DS9 can be used to obtain the counts in any pixel and thus can be used to determine an approximate count for the average sky or to check the number of counts detected in a star.

3.1.3 Gaia/SKYCAT

The Gaia (Graphical Astronomy and Image Analysis Tool) image software is an image viewer and analysis tool created by the ESO. It is a derivative of the ESO

SKYCAT software (Albrecht et al., 1997). It can be used as an image display facility with all the usual features as well as some particular astronomy ones (for instance photometry and defect patching) and it also can be used to access a a number of sky catalogues and as such can be used for astrometric calibration. For this research we used Gaia/SKYCAT to translate the x,y pixel coordinates used in the reduction to sky coordinates for publication, §3.9.1.

3.1.4 ISIS

ISIS is an image subtraction codeset designed to process CCD images using the image subtraction method developed by Alard and Lupton (1998). ISIS contains 3 sub-packages, and allows the user to build lightcurves of variable objects from a series of CCD images. The largest part of the ISIS process is the image subtraction. The three sub-packages are as follows:

Image registration:

The goal of image registration is to re-map each image onto the same pixel grid. The reference system, or common grid is usually one of the images. The output of this procedure is an image interpolated on the reference grid.

Image Subtraction:

This is the main program. A good reference image is made by stacking a number of the best images. Then the image subtraction code adjusts the reference image to the seeing of each individual image by means of a kernel convolution and performs the image subtraction. The essence of the image subtraction process is to find a convolution kernel that will optimally transform the reference image to match the target image. If R(x,y) is the reference image, and I(x,y) the target, then we require the kernel K(x,y), where

$$R(x, y) \otimes K(u, v) = I(x, y) \tag{3.1}$$

The kernel is decomposed as a linear combination of basis functions, B_i(u,v)

$$K(u,v) = \sum_i a_i \times B_i(u,v) \tag{3.2}$$

In ISIS, the basic functions are fixed-width Gaussians multiplied by polynomials.

The problem reduces to a linear least-squares, which is minimised to determine the coefficients a_i (which are allowed to vary spatially). The final output of the code is a subtracted image of the flux variation between the individual image and the reference frame.

Photometry:

This package creates photometric lightcurves from the subtracted images using a list of coordinates (determined from the stars in the reference image). The flux of the star is calculated using profile fitting photometry at a fixed position. For our survey we use the DAOPHOT package to determine these fixed stellar positions on the reference frame.

The image reduction used in this research is based on these steps and some of the code written for ISIS.

3.1.5 PySIS

The PySIS code is an image subtraction pipeline written by Dr Michael Albrow in PYTHON to assist with the reduction of data from microlensing events. Originally written to perform photometry on a single star, it has been modified to compile lightcurves of all stars whose coordinates are given. PySIS2 uses the ISIS code within a framework creating an easy to apply script. It incorporates all the sections outlined in §3.1.4 into one script which produces data files containing the lightcurves in flux, and separate files containing the reference flux as found from the reference frame. A new version of the PySIS software version 3.0 (Albrow et al., 2009) was used for the final data reduction in this thesis.

PySIS3 implements the image subtraction method as discussed but uses a numerical pixel grid (Bramich, 2008) rather than a sum of Gaussian functions to define the kernel. The pixel grid eliminates the requirement for any assumptions about the form of the kernel as it allows it to take any shape required. The method developed by Bramich (2008) retains the linearity of the least squares problem but is implemented with fewer parameters than the Alard and Lupton (1998) implementation used by the standard ISIS codeset.

3.2 Pre-processing

Prior to reduction of the images to lightcurves, every image in the dataset was corrected for bias and chip sensitivities. When the observations were taken, a number of calibration frames were acquired. The frames utilised in this reduction were the bias frames and the flat field frames. Bias frames are instantaneous images by the CCD camera with a positive offset imposed by the electronics so that pixels with zero detected photoelectrons always have a positive value. Flat field frames are images of an evenly illuminated surface: either the sky at dawn or an illuminated screen within the telescope dome. In this case the frames used were dome flats. The dome flats were taken with the appropriate filter, in this case the R filter (§2.1.2), and allow the variations in sensitivity across the CCD mosaic to be determined and therefore corrected for. Pre-processing involves subtraction of the bias frame and division by the flat field frames.

The raw data used for this research are mosaic images composed of 8 CCD frames, (Chapter 2). Specific parameters and commands are needed in order to perform operations on all chips of the mosaic at one time. The IRAF group at the National Optical Astronomy Observatory (NOAO) created the ESOWFI and MSCRED packages for users with data from the ESO WFI camera. ESOWFI loads both packages and sets up the correct default parameters for WFI data. The ESOHDR

command was performed on every image in the dataset to ensure that the correct keywords were present in the header files before the MSCRED package was used to pre-process the images. The MSCRED command package performs standard IRAF commands on the entire mosaic frame without the need to split the frames and can also be used to divide the images or join them back into the mosaic.

All the flat field and science frames had the appropriate bias frame subtracted off them. Each night of observations has 4 or 5 bias frames associated with them. These frames were median combined using ZEROCOMBINE from the MSCRED package to create a master bias frame for each night of observations. This command is specifically designed to work with bias frames where there is normally little variation across the frame. Providing that the bias frames show little variation with time an overall master bias can be created. The observations used in this research in some cases showed some variation with time and the decision was made to use a master frame where there was minimal variation and nightly bias frames otherwise.

Following subtraction of the bias, the images were divided by the appropriate flat field image. Like the bias frames there are multiple flat field images for each night of observation and these were also median combined, this time using FLATCOMBINE from the MSCRED package, to create master flats for each night. The subtraction of the bias from the science frames and the division of the science frames by the flat field frames was accomplished by use of the MSCRED command CCDPROC. CCDPROC subtracts the appropriate bias frame and then divides the resultant image by the appropriate flat field.

After the mosaic images had been pre-processed they were divided into individual FITS files for each CCD chip. The programs used for the reduction were not designed to work on mosaic images. The MSCRED command MSCSPLIT was used to split the mosaics. This led to eight file sets with extensions 1 through 8.

3.3 Image registration

The first step in the image subtraction process is the registration of the images. To perform image subtraction properly the stars must fall on the same set of pixels in each frame. For ISIS this requires a translation to the exact coordinates in each frame, and involves interpolation of the images. The PySIS code used in this research only requires that the images are registered to the nearest pixel and thus only integral-pixel shifts are performed (no interpolation).

The best-seeing image is determined and this is used as the registration template. A selection of bright stars with their pixel locations is found for every image and the translation between the image and the registration frame is determined. The translation is then used to rewrite the image on the new coordinate frame. The registration process is extremely important and must be performed correctly before any other step can be completed.

In the images used in this research, problems were found with the registration of the large, rectangular images. The number of stars in the large frames caused confusion in the registration process as different stars were identified in similar patterns in different frames. This caused unpredictable large coordinate translations rather than the small visibly detectable shifts required. Also, a simple linear x-y shift proved insufficient to register an entire 2k by 8k frame due to small relative distortions and rotations. To solve this problem the 2k by 4k images were sectioned into eight 1k by 1k sub-images. This resulted in 64 file sets (eight sections of eight CCD chips) given prefixes A through H. These smaller images had no problems with registration.

3.4 Reference frame creation

Using all the pre-processed and registered image files PySIS3 determines the seeing for each of these in order to select the best seeing images for compilation into a

photometric reference image. The cutoff point of images is when the seeing becomes significantly different from that of the best image. PySIS normally seeks a list of up to ten images to compile into a reference frame, provided that they all fall within the delta-seeing cutoff. The reference frame for this dataset was chosen to be a stack of three of the best seeing images in the dataset. All these images had the best seeing over all eight chips and were all from the 2005 dataset. Originally it was attempted to create a reference frame with the five best-seeing images, but small field distortions between the 2002 and 2005 data caused distortion in the reference frame and so the decision was made to utilise only three frames in creating the reference.

Following compilation of the reference frame, the DAOPHOT code in PyRAF (§3.4.1) was used to find stellar positions and fluxes on the reference image.

3.4.1 DAOPHOT

DAOPHOT is a stellar photometry package designed by Stetson (1987) to deal with crowded fields. The package performs various tasks including finding objects, aperture photometry, obtaining the point-spread function, and profile-fitting photometry. Profile fitting in crowded regions is performed iteratively, which improves the accuracy of the photometry.

We used DAOPHOT to determine positions and fluxes for stars on our reference frames following the guidelines in the IRAF PSF documentation as follows. The DAOFIND command was utilised to find all stars above a threshold limit of 3σ. This resulted in between seven and thirteen thousand sets of star coordinates for each of the 64 sub-images. The command PHOT was then run on these star positions to obtain relative magnitudes for the stars. Given the relative magnitudes, several stars were selected using the command PSTSELECT for creation of the point spread function (PSF). Each of these stars (around 25 per sub image) was examined by eye for distortions and saturation. In some cases all the selected stars had to

be eliminated from the list upon examination due to saturation (The resulting PSF from using saturated stars appeared quite distorted with spikes towards the top and bottom as saturated stars bleed along columns; (Figure 3.1)). To solve this problem a larger number of stars are selected and those at the end of the list were used rather than those at the beginning. This allowed a set of bright but unsaturated stars to be selected for use in creating the PSF. The resulting list of suitable stars was run through the command PSF to create an initial PSF. This PSF was of Gaussian form with numerical residuals. It is possible to allow the PSF to vary over the chip, however we found that a constant PSF was suitable. In some cases the selection of stars resulted in a PSF image which would not converge despite the examined stars appearing good. To resolve this, PSTSELECT was rerun to obtain more stars for inclusion.

Following the creation of the initial PSF, the neighbouring stars of each of the PSTSELECT stars were determined using NSTAR and then these and the PSF stars were subtracted from the reference frame using SUBTRACT. The resultant image was examined to find any neighbouring stars that were concealed by the PSF stars when using DAOFIND and NSTAR. Undetected stars were added to the list of stars with coordinates so that all neighbouring stars were subtracted to leave an isolated star for secondary PSF construction. When all the neighbours had been detected, the list of neighbours and PSF stars created by NSTAR was modified to remove the PSF stars from the list. When this secondary list was run through SUBTRACT, an image with isolated PSF stars was created and from this image a new PSF was constructed. The PSF from this step should be a good PSF for use in calculating reference fluxes for all the detected stars. In several of the sub-images, the PSF was not a good fit so the PSTSELECT stars were re-examined for shape and for missed neighbours. These steps were repeated until a good PSF is found (Fig 3.1).

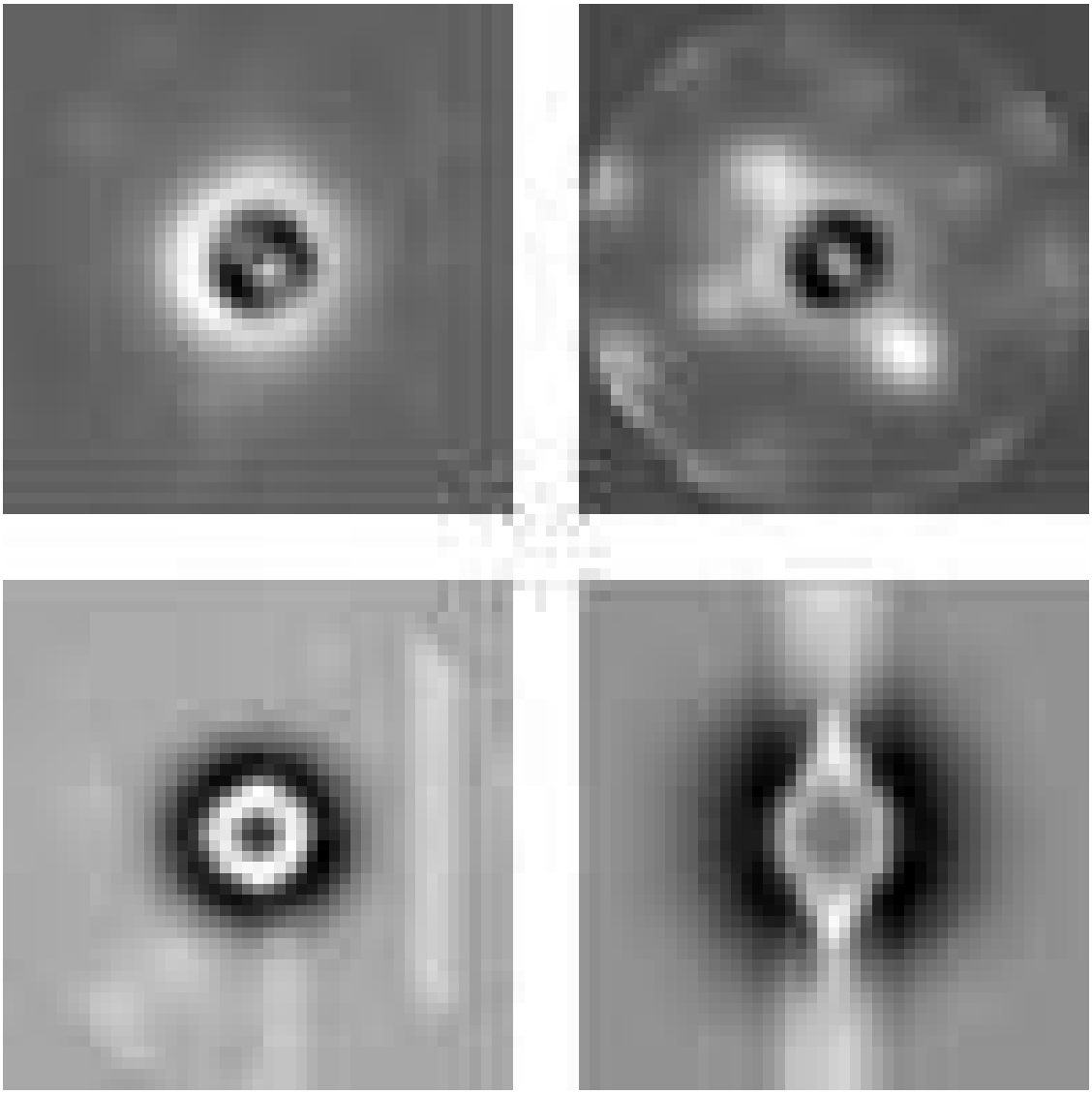


Figure 3.1: **Point spread function** Images of the non-Gaussian components of the point spread functions as obtained in analysis of the reference frames resulting from: a good PSF (top left), a PSF with neighbouring stars (top right), a PSF with strong saturation nearby (bottom left) and a PSF created with saturated stars.

With a good PSF image and an overall list of stars with relative magnitudes the command ALLSTAR was run on the original frame with the final PSF image. ALLSTAR creates an output with the coordinates, magnitudes and error in magnitudes of all the stars of sufficient number of pixels and magnitude to be distinguished from the background.

In this way, DAOPHOT was used to perform photometry on the reference frame to determine magnitudes in relation to a zero point, and stellar positions of all identifiable stars. These stellar positions were then fed into the PySIS code (§3.6) to create lightcurves.

3.5 Image subtraction

The deep reference template was created from three of the images with the best seeing (§3.4). This template was mapped to and subtracted from each individual image. The reference image and registered frame are the input for the subtraction routine. A saturation mask of the reference and target images is compiled and the sky background removed from each. The kernel (of variable size - scaled to match seeing variations) was computed (Bramich, 2008; Albrow et al., 2009) and the reference frame convolved with the kernel. This convolved reference is then subtracted off the image frame. Finally the difference image is divided by the integral of the kernel, to place it on the same effective exposure scale as the reference image.

The scaling of the kernel size was tested with several different seeing expansions. The best result - fewest residuals - was found when the kernel size was set to be 1.5 times the seeing FWHM. This worked well for good and average seeing images. The images with bad seeing still had problems, but any additional modification made things worse overall. The kernel was allowed to vary across the image, each kernel pixel having a bi-quadratic spatial variation. The outputs from the subtraction routine are the subtracted images, normalised subtracted images,

the saturation masks, kernel and kernel tables, and the model images. A sample subtracted image can be seen in Figure 3.2.

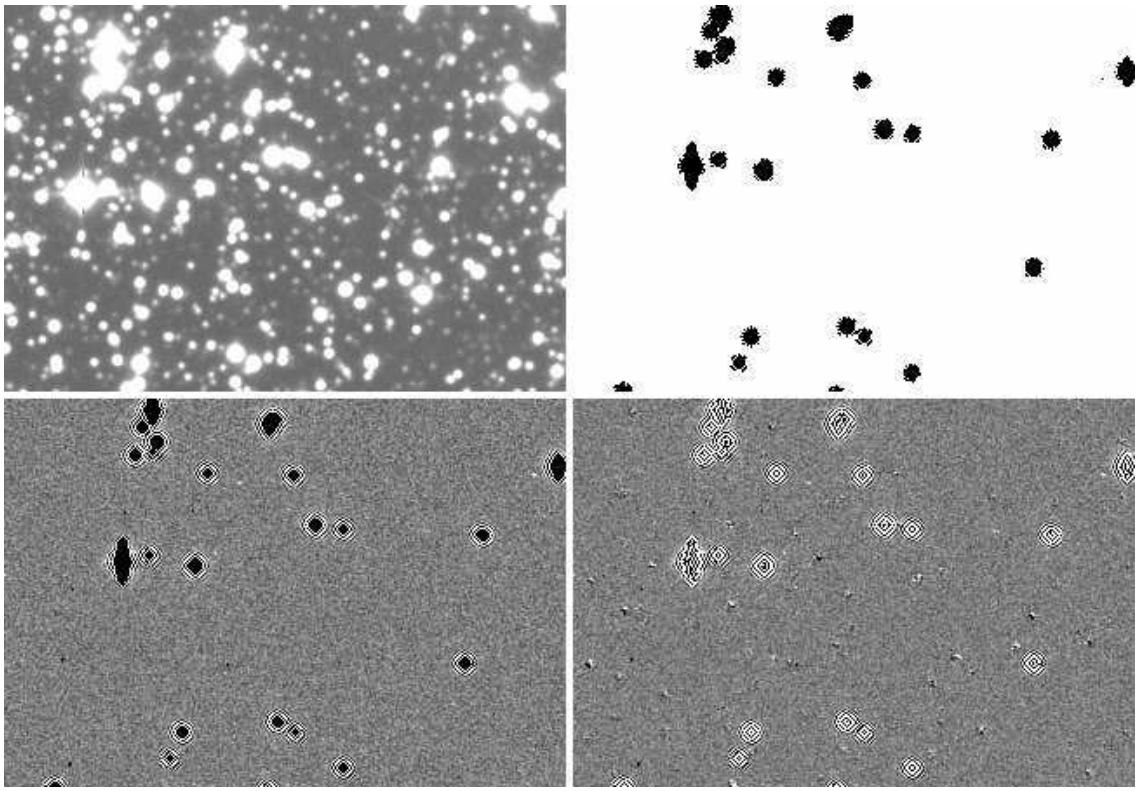


Figure 3.2: Reduction of images goes through several stages. Top left the original registered image, top right the saturation mask, bottom right the convolved and subtracted image, bottom left normalised version of same i.e. the difference image in Poisson noise units. Residuals left by the subtraction can be seen around the saturated stars. Each frame of the image is 76.2" by 52.7"

3.6 Photometry

PSF photometry of the difference images was performed for all detected stars. The photometry routine has inputs of the subtracted image, the coordinate list produced by DAOphot and the kernel table produced by the subtraction routine. Also, we need the PSF of the reference image. We used the ISIS Bphot routine to create this PSF. For each star the photometry routine computes the mean sky from 50

x 50 pixel image sections centred on the target. The reference PSF is convolved with the kernel and then mapped onto a pixel grid at the target coordinates. The difference flux for each star is computed by optimal fitting of the PSF at the target coordinates. Results from this are written to the lightcurve file. The photometry is also performed on the reference image (without using kernel convolution) to obtain reference fluxes for the stars detected by DAOphot. The DAOphot routines also produce reference values - in magnitudes. Using the zero-point for the magnitudes found by DAOphot the PySIS reference fluxes were transformed to a magnitude scale and are compared with the DAOphot reference magnitudes in Figure 3.3.

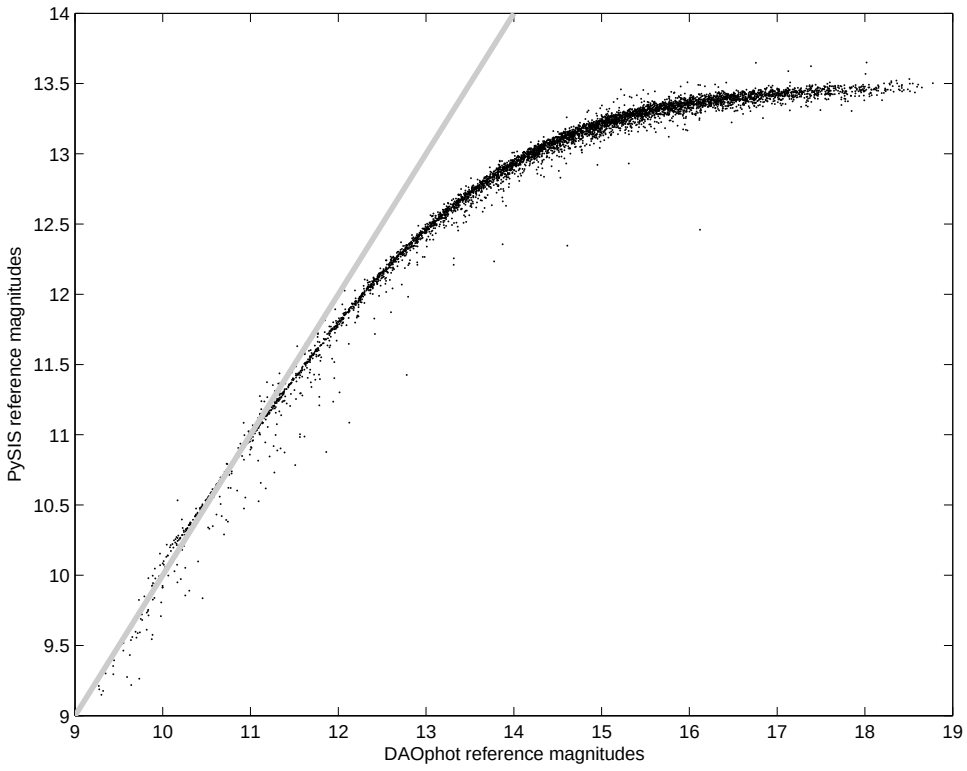


Figure 3.3: Comparison of reference magnitudes obtained via PySIS reduction and DAOphot computation overplotted with a line of slope one. The reference magnitudes are a good match at the brighter magnitudes. The values differ increasingly to fainter magnitudes which is expected from the way each program treats blends.

This is a good determination of whether the photometry is being performed correctly. The reference magnitudes from PySIS are a good match to those obtained from DAOphot at the brighter magnitudes. The increasing difference in the values from here is expected and caused by the difference in treatment of blends by the two programs, i.e. DAOphot ALLSTAR fits to multiple stars simultaneously while the PySIS fluxes for fainter stars are contaminated by neighbours and underestimation of the background. We use the DAOPHOT reference magnitudes in later work.

Photometry of the image frames is performed on each frame separately and within the routine a loop ensures that the photometry is performed on each set of star coordinates. Each lightcurve was assigned a filename based on the star number designated by DAOphot.

3.7 Implementation and results

The application of the PySIS code results in lightcurve files containing time of observation, flux and uncertainty for each epoch, plus secondary reference files containing the flux and uncertainty from the reference frame. The resulting precision for a 500 by 500 pixel section of one of the chips from the data acquired by the PLANET collaboration in 2002 (§2.2) can be seen in Figure 3.4. In comparison to the Poisson noise limits for the particular magnitude, the data can be seen to have a high precision and is clearly suitable for analysis for planetary transits. This shows that the reduction software and method detailed in this chapter is suitable for use in the search for extrasolar planetary transits.

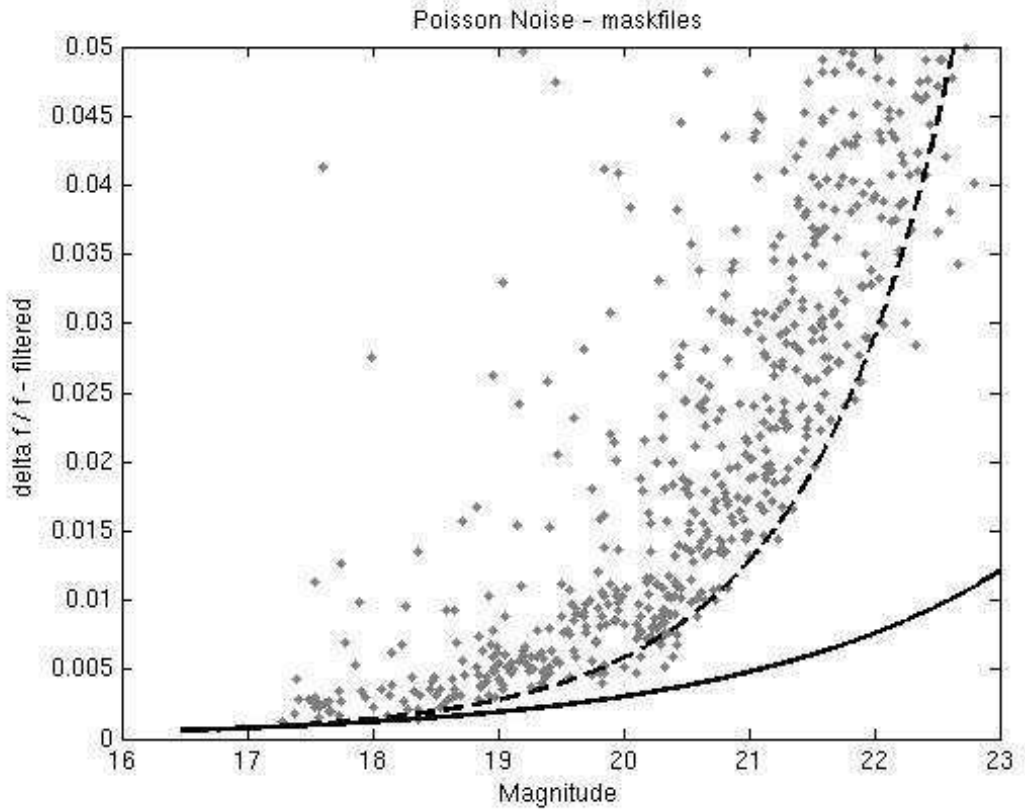


Figure 3.4: The precision of data analysed by ISIS and DAOphot including the Poisson noise limits with (–) and without (-) sky background. The magnitudes are uncalibrated instrumental magnitudes.

The method detailed above was, in an earlier form, first used to re-reduce the data collected in 2002 to determine suitability of both the data and the reduction method. The results from this reduction were used in the proposal detailed in §2.3 to request an new set of observations in the same field. The new observations collected in mid 2005 and released at the end of 2005 were originally slated to be reduced to lightcurves by Cristina Afonso at the Max-Planck Institut fuer Astronomie in Heidelberg. Due to large numbers of computer problems this did not occur and in March 2007 the observations became available to the public. The observations were then requested to be sent from the ESO archive to the University of Canterbury. They arrived in June and, after problems with the data computer were remedied, were ready to be reduced in early August of 2007.

After pre-processing as detailed in §3.2 the decision was made to also request the 2002 data from the archive rather than attempting to extract it from a failed hard-drive. This was accomplished by means of a high-speed ftp link newly installed at the University of Canterbury and arrived within a few days of the request.

These images were also pre-processed and then the complete data set of 267 images was reduced to lightcurves using the methods detailed in §3.2 - §3.6 using the update PySIS3 code. Not all the images were suitable for reduction but between 261 and 265 data points were obtained for each star, except those located near to the edge of CCD frames which, in a number of cases, lost some points due to variations in pointing.

Due to the time taken by the convolution of each image in the subtraction process use was made of the University of Canterbury Super Computer to complete the subtraction and photometry routines in parallel. Each sub-image subtraction takes between one half to two hours depending on the seeing of the image (the kernel size is scaled by the seeing). The total computation time for 64 sets of sub-images (1k by 1k sections) containing 267 images each was estimated at approximately 530 days if run in series on a single processor. Use of the supercomputer allowed the

computation time to be drastically reduced to approximately 6 weeks.

Our complete reduction resulted in approximately 500,000 lightcurves, all recorded as individual files, containing the Julian data of observation, δ flux, and error in δ flux. In order to use these lightcurves, the individual files are combined into three matrices - one each for time, δ flux and its uncertainty.

3.8 Lightcurves

Between six and eleven thousand lightcurves were produced from each of the the 64 subsections of the images. A large fraction of these has sufficient precision for variability searches, Figure 3.5 shows the RMS of a subset of lightcurves. Before the search algorithms are run a number of scripts were run on the data to make it ready for the searches. The initial steps are to divide the data points by the appropriate reference magnitude to obtain δflux/flux and then to sort the data so that the lightcurves are all in temporal order.

With the lightcurves extracted from the images it is important to remove any extraneous values and any systematics. The lightcurves in this dataset are less likely to have systematics than most datasets due to the method of reduction, but as discussed in §2.2.2, systematics can have a large effect on the planet yield and so it is important to remove as much as possible. Image subtraction uses a single base image for the subtraction and so systematics present in this image will be removed from all images. Different systematics may appear in multiple other frames and thus the dataset should be treated with a detrending method. The following sections discuss the cleaning and detrending steps implemented on this dataset before the searches were performed.

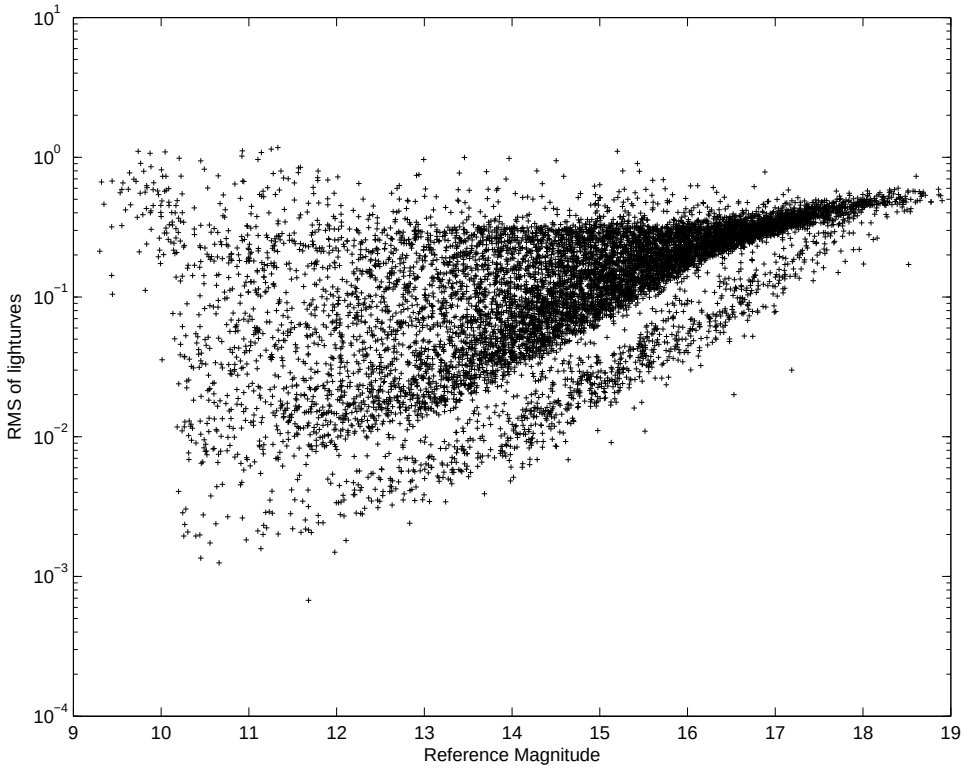


Figure 3.5: **Sample dataset RMS** RMS of a sample dataset before any filtering or detrending as plotted against uncalibrated reference magnitudes. The bifurcation of the upper data appears as a result of trending effects in the reduction. The lower outliers are a results of stars with small numbers of points due to the reduction process.

3.8.1 Cleaning

When the lightcurves are first reduced and before they are fed into the search algorithms, the lightcurves require some cleaning or filtering. The simple filtering we employed involved the removal of extraneous values individually by means of a 4-sigma level cut. The points were removed individually, and the RMS recomputed at each iteration. This is more robust than removing all points outside a threshold value at once. A maximum of 15 points were removed from each lightcurve. A sample lightcurve before and after cleaning is shown in Figure 3.6.

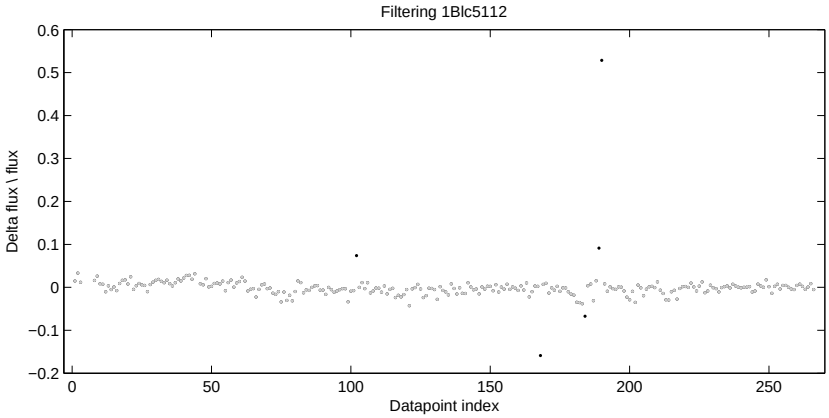


Figure 3.6: **Filtering effects on a single lightcurve**, black • indicates the removed datapoints, the grey •, are the post-filtered datapoints.

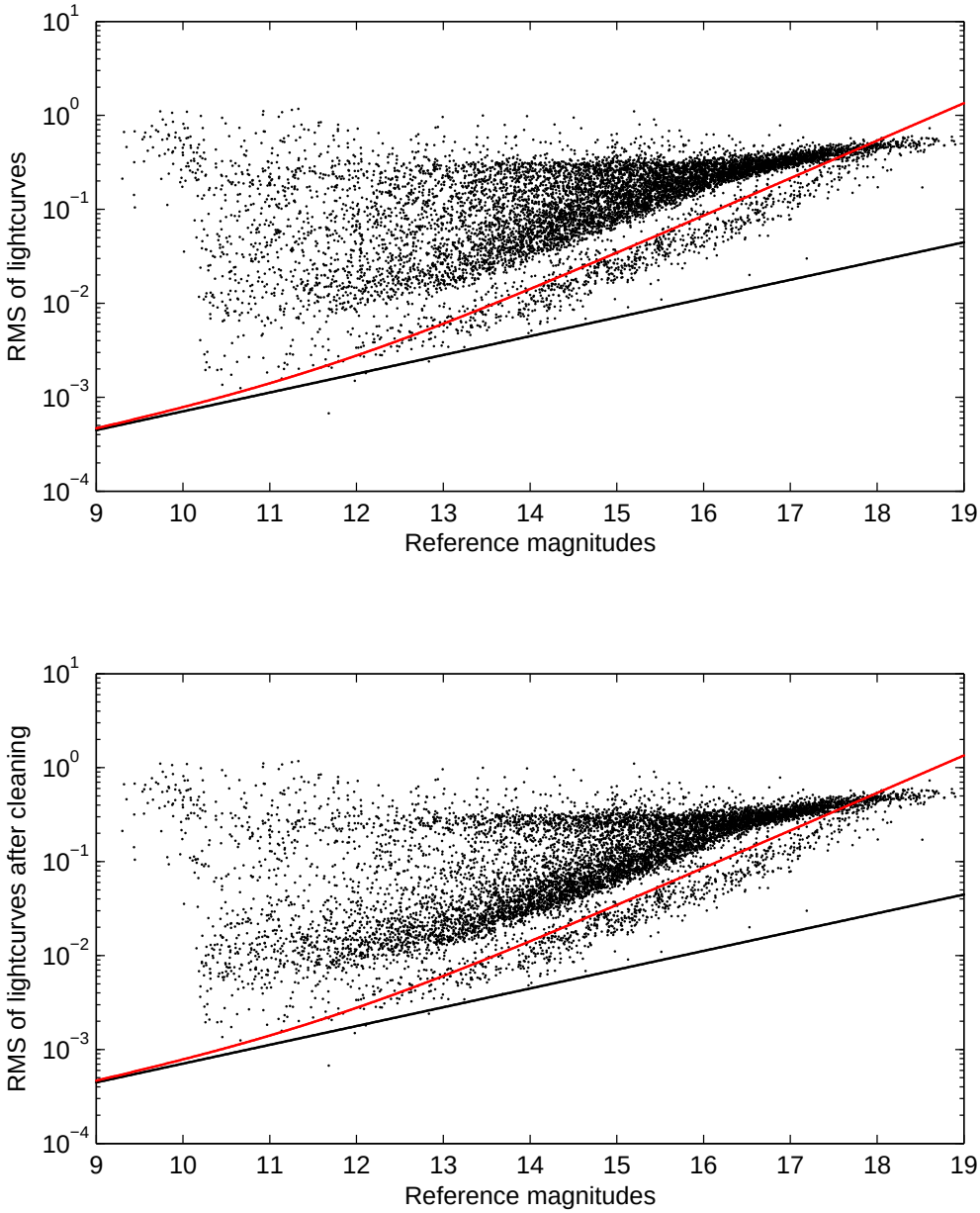


Figure 3.7: **Filtering of a sample dataset** The results from running the filtering script on a subset of the lightcurves, the upper panel shows the RMS of the uncleaned lightcurves, and the bottom panel the results from cleaning. Also plotted are Poisson noise with (upper line) and without (lower line) sky and readout noise.

3.8.2 Detrending

As previously discussed in §1.4.7.3, large photometric datasets have trends in the data resulting from systematic noise (red noise). Combined with the white noise from the data the systematics can have a large effect on the extrasolar planetary yield from a transit survey. Systematics are caused by such things as atmospheric extinction changes over time, detector efficiency or changes in the points spread function (PSF) over the detector. PSF changes are particularly relevant for this work as the large size of the detector chips means more chance of variations with position. The atmospheric extinction varies with position in the sky and will change depending on the time of night the observations were taken. Depending on the coordinates of the observation field this may have more effect on some datasets than others. The detector efficiency is generally assumed to be constant over short time scales but some variations may be present.

Using image subtraction for the reduction removes a large portion of the systematics, but efforts to detrend the dataset must still be made. There are a number of different detrending methods available. In general detrending involves the use of a template set of stars to determine any trends in the dataset, the results from the template set are then applied to the full dataset. There are two detrending algorithms that are used most commonly for transit surveys. They are designed for large datasets or wide-field surveys and both were designed in response to the lack of planets discovered despite good predictions in a number of transit surveys.

Designed by Tamuz et al. (2005), SysRem is used to correct systematics in a large dataset of photometric lightcurves. Systematic effects are found and corrected providing the effects are noticeable in a large number of the lightcurves in the dataset. SysRem was designed such that it would determine any sort of systematic effect without any prior knowledge of the types of systematics it was correcting. SysRem calculates two sets of coefficients which best represent the systematics that are being corrected. This is based on the 'criss-cross' iterative process proposed

by Gabriel and Zamir (1979). That process is very similar to the SysRem algorithm but was used for climate statistics and chemometrics, mostly to discuss convergence rather than remove trends.

SysRem was originally developed to correct colour-dependent atmospheric extinction and we discuss here this implementation. A set of N lightcurves (with M measurements) has residuals, r_{ij}, defined as stellar magnitudes with individual average magnitudes subtracted. The extinction coefficient, c_i, is defined as the slope of the best linear fit of the residuals, r_{ij}; j = 1,...,M; i = 1,...,N as a function of the airmass, a_j; j = 1,...,M. The best c_i is found when

$$S^2 = \sum_j \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{3.3}$$

is minimised. σ_{ij} is the uncertainty of the measurement of star i in image j. Once c_i is found it can be held constant to determine a set of a_j which will best minimise Equation (3.3). This a_j may have no relation to the actual airmass but is "the most suitable airmass of each image" (Tamuz et al., 2005). This method can be generalized for any effect which varies linearly with a stellar property and the process can be iterated for any number of systematic effects.

The other major algorithm is the trend filtering algorithm (TFA) proposed by (Kovács et al., 2005) and this is what we have implemented in this work. TFA was initially designed for use in wide-field variability surveys and is suitable for transit surveys. It was tested on data from the HATNet project (§1.4.7.2) and is designed to be applied to photometric time-series data. The TFA corrects effects from instruments and also those that may arise from the data reduction process. The primary purpose of TFA is to remove trends but it may also be used in reconstructing periodic signals in a dataset. This purpose is not necessary for our application and will not be further discussed.

TFA assumes that there are a large number of stars spread over the CCD and

that a large number of these stars will be non-varying. In the case of this dataset a large number of stars are assumed to have close to zero values of δflux/flux (non-varying). Following processing by the TFA these stars should be dominated by white noise.

It is assumed that all the lightcurves have observations at each time (i = 1,2,...,N). For our data, this involved reshuffling the lightcurve matrix to ensure that each time corresponded to a single column. The filter:

$$F(i) = \sum_{j=1}^{M} c_j X_j(i) \tag{3.4}$$

is built up from the time-series, X_j(i), of a set of template stars (j = 1,2,...,M). Template stars are selected such that they cover the coordinate space, have a large number of datapoints (missing points are filled with an average), low standard deviation and a good level of brightness.

The coefficients of detrending are then determined from minimizing the expression:

$$\mathcal{D} = \sum_{i=1}^{N} [Y(i) - A(i) - F(i)] \tag{3.5}$$

$$\text{where} \quad A(i) = \langle Y \rangle \equiv N^{-1} \sum_{k=1}^{N} Y(k) = \text{constant}$$

where Y is the time series being filtered and A the best estimate of the detrended time-series assuming the time-series is constant and thus dominated by noise and systematics and not containing any periodic trend. If the lightcurve does contain a periodic trend, detrending in this manner can have a detrimental effect on detectability of the periodic signal. In general the amplitude of the variation is lowered. Effects have been noted in the variable stars found in this research (see Figure 3.8) where the amplitude has been significantly lowered to where the lightcurve is barely periodic. Significant detrimental effects were not seen in the transit lightcurves

(see Figures 5.17 to 5.37 for examples). The transit (§5.4) and variable searches (§6.2) were run on both the detrended and the un-detrended data.

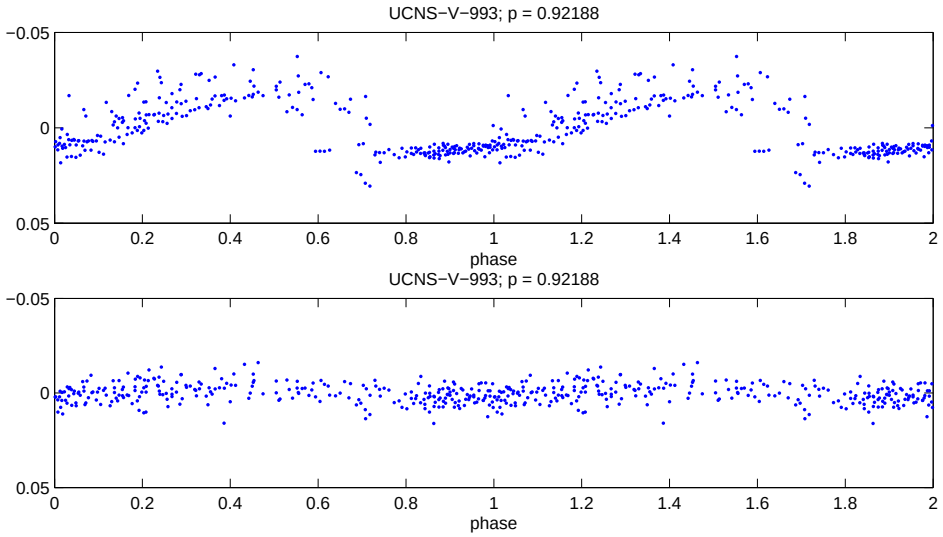


Figure 3.8: **Detrending of a variable star** The effects of detrending on variable stars. The upper panel shows a variable star detected by the search when run on the filtered dataset. The lower panel shows the same star after detrending. The amplitude in the detrended lightcurve is significantly lower than the amplitude of the un-detrended lightcurve.

In application, a normal matrix, $g_{j,k}$, is computed from the template lightcurves

$$g_{j,k} = \sum_{i=1}^{N} X_j(i) X_k(i); \qquad j,k = 1,2,\ldots M \tag{3.6}$$

and its inverse G computed. The normal matrix must be non-singular or as close to as possible. Missing datapoints or very small datapoints can cause problems and the template stars must be selected so that this does not occur. These steps are computed once for the dataset as it is based solely on the template stars. Following, and for each lightcurve in the dataset, a scalar product:

$$h_j = \sum_{i=1}^{N} \tilde{Y}(i) X_j(i) \tag{3.7}$$

$$\text{where} \quad \tilde{Y}(i) \equiv Y(i) - A(i)$$

is computed. The solution of the coefficients for the star is determined:

$$c_j = \sum_{k=1}^{M} G_{j,k} h_k \tag{3.8}$$
$$\text{where} \quad G_{j,k} = \text{inverse of } g_{j,k}$$

and then the corrected lightcurve constructed.

$$\tilde{Y}(i) = Y(i) - \sum_{k=1}^{M} c_j X_j(i) \tag{3.9}$$

In our dataset a sample of 100 template stars was chosen to sample each subset. The template stars were chosen from those with a large number of data points and with magnitudes in the mid-ranges of the survey (12 - 17). The coordinates of the template stars were plotted over the coordinates of all the stars to check for good spatial coverage. Following template selection any missing data points in the template lightcurves are filled with the mean of the lightcurve, this should be approximately zero. Using the template stars the normal matrix was constructed and its inverse found, Equation (3.6). For each star the product of the target and template time-series was found, Equation (3.7). Using these two values the coefficient c_j, was computed for each lightcurve and then used to correct the original data.

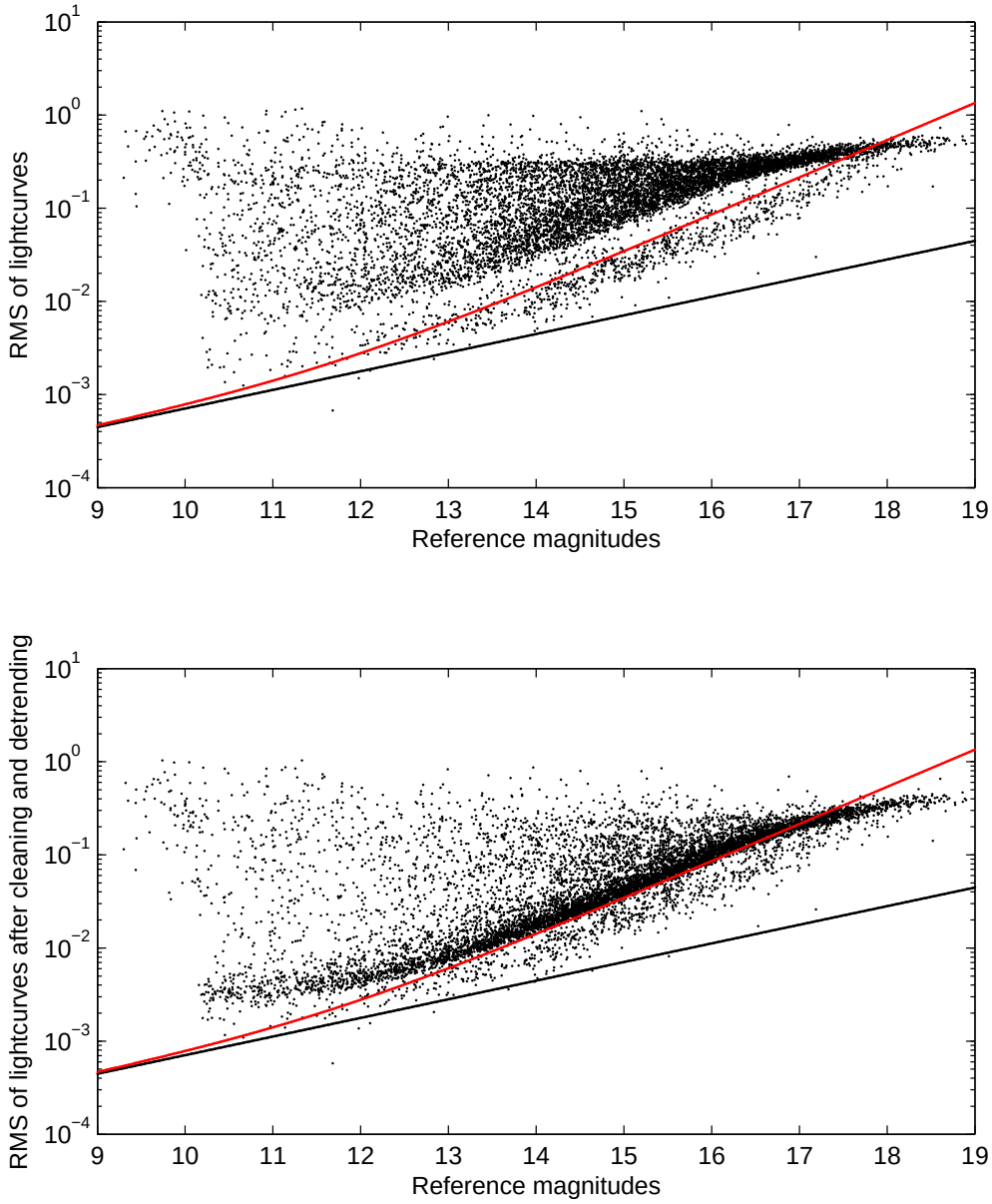


Figure 3.9: **Detrending of a sample dataset** The results from running the detrending algorithm on a subset of the lightcurves, the upper panel shows the RMS of the uncleaned lightcurves, and the bottom panel the results from detrending. Also plotted are Poisson noise with (upper line) and without (lower line) sky and readout noise.

3.8.2.1 Poisson limits

The Poisson noise is calculated with the following equation:

$$Poisson\ noise = \sqrt{\frac{npixels\ read\_out\_noise^2}{gain^2} + \frac{npixels\ sky\_level}{gain} + \frac{flux}{gain}} \quad (3.10)$$

where npixels is the effective number of pixels used in the optimal PSF-fitting photometry. This number depends on the PSF width and is different for each image. From Albrow et al. (2009),

$$npixels = \frac{\pi(FWHM)^2}{2ln2} \frac{1 - e^{\frac{-4ln2R^2}{(FWHM)^2}}}{1 - e^{\frac{-8ln2R^2}{(FWHM)^2}}} \quad (3.11)$$

for a Gaussian PSF truncated at radius R. We adopted an average FWHM = 6.5 and hence npixels $\approx$ 140.

3.9 Calibration

3.9.1 Coordinate calibration

Using the astrometric capabilities of Gaia (3.1.3) the x,y pixel coordinates of the reference frames have been transformed to sky coordinates. The x,y coordinates of the stars used in the reduction are then transformed by use of MATLAB transformation matrices. We display our reference frame in Gaia and an image at our coordinates from the Digitized Sky Survey[1] (DSS) using the Aladin applet (Bonnarel et al., 2000). The coordinates from a United States Naval Observatory (USNO) catalogue (Monet et al., 2003) are then overplotted in Aladin. By identifying stars on the DSS image which have coordinates from the USNO catalogue and are unsatu-

[1]This research has made use of the Digitized Sky Surveys, produced at the Space Telescope Science Institute under U.S. Government grant NAG W-2166. The images of these surveys are based on photographic data obtained using the Oschin Schmidt Telescope on Palomar Mountain and the UK Schmidt Telescope. The plates were processed into the present compressed digital form with the permission of these institutions.

rated in our reference image, an astrometric calibration is performed in Gaia. The known coordinates are then used in MATLAB to create a transformation matrix which is then applied to the x,y coordinates of all the stars. This has resulted in known right ascension and declination (in J2000) for all the stars in the field. These coordinates are then used in the transit candidate and variable star identifications.

3.9.2 Magnitude calibration

Similarly to the coordinate calibration (§3.9.1), approximately 150 stars were identified in the reference frame from the USNO catalogue (Monet et al., 2003). The magnitudes were recorded and compared and an offest for the magnitude calibration determined. Using this we were able to make a preliminary calibration of the instrumental magnitudes. The average offset for the R magnitudes over the field as compared to the USNO B1.0 R1 magnitudes is -5.4027 and all our instrumental magnitudes have been corrected for this. The corrections for B and I were made in similar fashion. We note that no colour-dependent terms have yet been applied and that our "calibrated" magnitudes are solely corrections to put them on the USNO B1.0 scale.

3.9.3 Colour-magnitude diagrams

The dataset contains three each of B and I images as well as the primary R images. These were pre-processed using the method discussed in §3.2 and subsectioned as for the R images. Using the reference frames from the R images as astrometric registration images, the B and I images were registered using the PySIS code. Once registered the B and I sub-images were combined into single frames using the methods used to create the R reference frames. As the images were now registered to the same coordinates as the R images we used the coordinates and star ids from the R images to obtain reference magnitudes by way of DAOPHOT for the B and I com-

posite images. The result is reference instrumental magnitudes for each star in B, R and I. These were calibrated (§3.9.1) and then used to create colour-magnitude diagrams upon which variable stars and transit candidates can be superimposed to determine position relative to the main sequence. Figure 3.10 shows the colour magnitude diagrams from each sub-image, and Figure 3.11 shows the same diagrams with a log density scale. The main sequence is strongly defined (greatest density in Figure 3.11) indicating that the field is mostly at the same distance (in the Norma Spiral Arm). At the top of the main sequence we can see a bifurcation to redder values which may indicate observations of a secondary more distant giant population. The blue spur found at the top of the main sequence is caused by stars saturated in only the R filter causing them to appear bluer. The colour-magnitude diagrams also show many field stars. These diagrams are reproduced in higher magnification in Chapter 6.

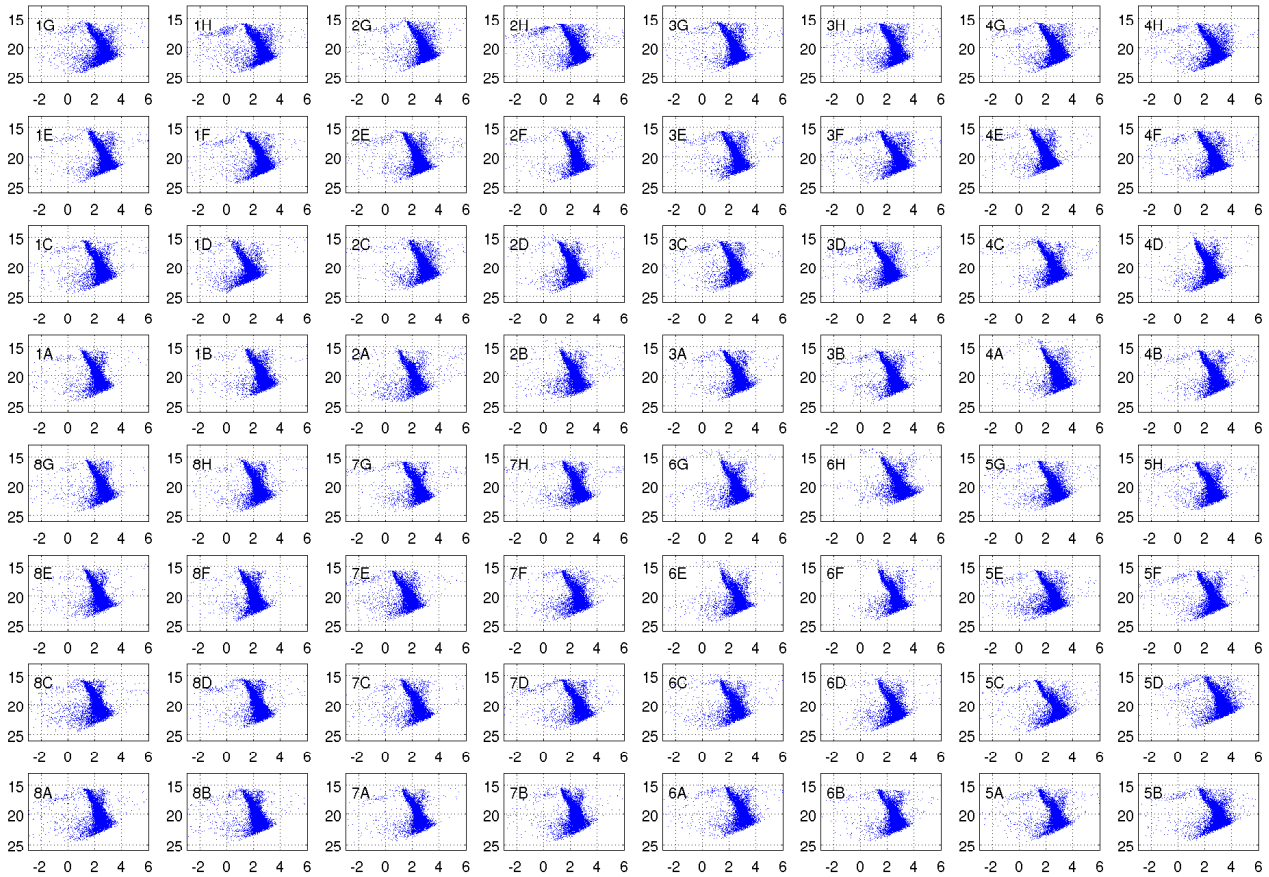


Figure 3.10: **Colour-magnitude diagrams** Each panel shows the colour-magnitude diagram (R vs B-R) for the stars of a sub-image with calibrated magnitudes. The diagrams are arranged in the same manner as the 1k by 1k sub-images appear in the mosaic.

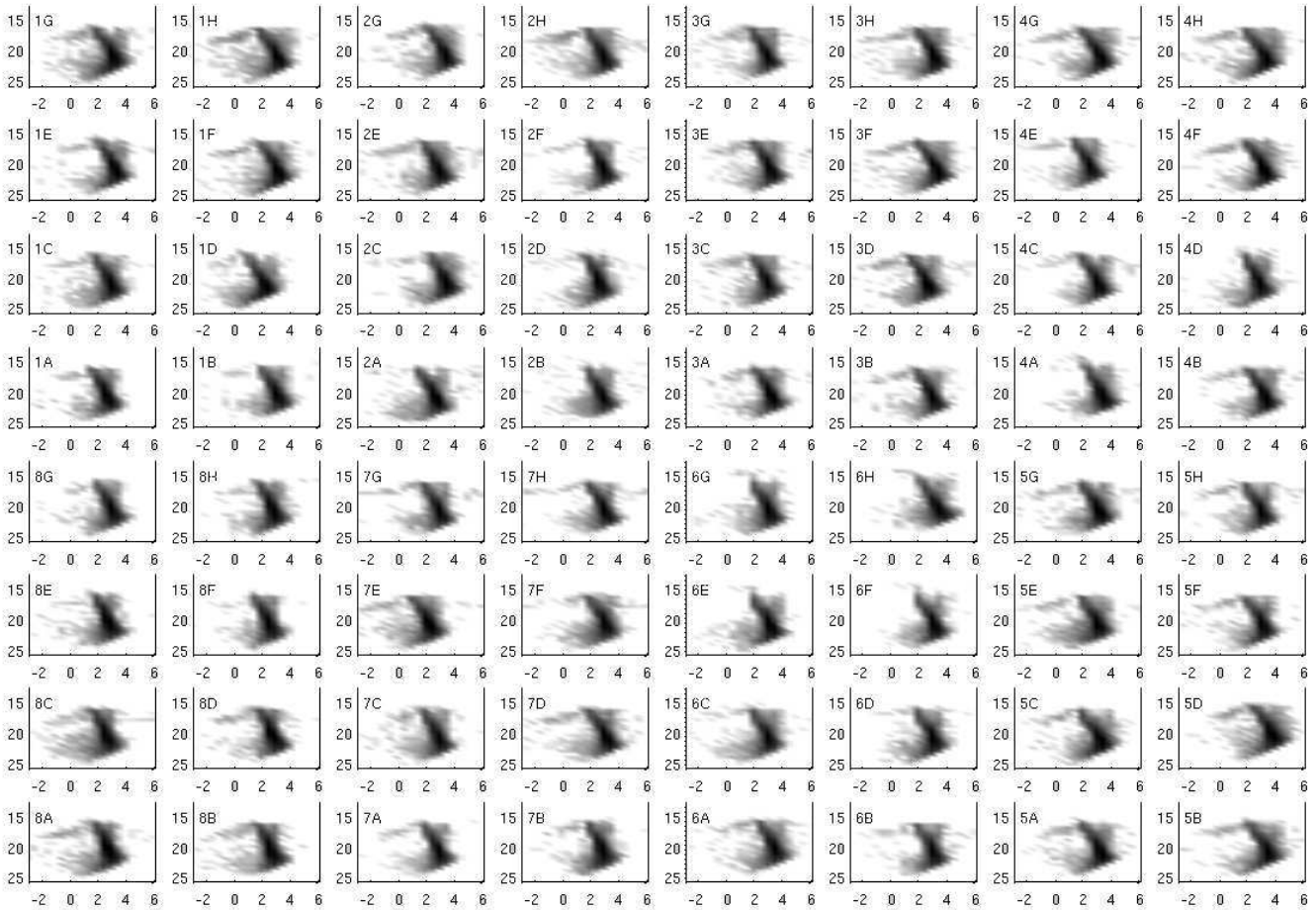


Figure 3.11: **Colour-magnitude diagrams (density)** Each panel shows the colour-magnitude diagram (R vs B-R) for the stars of a sub-image with calibrated magnitudes. The diagrams are arranged in the same manner as the 1k by 1k sub-images appear in the mosaic. The plots are shown with a log-density scale where greater density is shown darker.

Chapter 4

Search algorithms

To detect periodic variations in a large number of stars we make use of search algorithms. For each lightcurve the algorithms determine the best fit period and a measure of how significant the fit is, then a threshold can be determined whereby above (or below), the lightcurve is likely to have a periodic signal. Discussed here are algorithms for both transit searches and variable stars.

4.1 Transit identification algorithms

While the aim of all methods of transit detection is the same - identifying the parameters of a repeating dip in the intensity of a lightcurve - the base algorithms used to obtain a best parameter set for a lightcurve containing a possible transit are quite varied.

A comparison of the most common methods by Tingley (2003a), by application to synthesised lightcurves, was the starting point for an investigation into the suitability of algorithms currently in use for our data.

Correlation is the most basic statistical tool able to measure how well two datasets resemble each other. It is not designed for transit identification but is mentioned here as reference point for the transit detection methods. The basic

definition of the linear correlation coefficient is shown in (4.1), where D and S are the lightcurve and the trial transit signal, with $\overline{D}$ and $\overline{S}$ being their mean values.

$$T = \left(\frac{\sum_{n=1}^{N} \left(D_n - \overline{D}\right)\left(S_n - \overline{S}\right)}{\left[\sum_{n=1}^{N} \left(D_n - \overline{D}\right)^2 \sum_{n=1}^{N} \left(S_n - \overline{S}\right)^2\right]^{\frac{1}{2}}} \right) \tag{4.1}$$

4.1.1 BLS

The most commonly used search algorithm is the traditional box-fitting by least squares method as described by Kovács et al. (2002). This assumes that the lightcurve will take only one of two values, H outside the transit and L within.
A measure of the box signal fit, χ^2, to the lightcurve is calculated:

$$\chi^2 = \sum_{out} w_n (D_n - H)^2 + \sum_{in} w_n (D_n - L)^2 \tag{4.2}$$

which expands to:

$$\chi^2 = \sum_{all} w_n D_n^2 + H^2 \sum_{out} w_n - 2H \sum_{out} w_n D_n + L^2 \sum_{in} w_n - 2L \sum_{in} w_n D_n \tag{4.3}$$

$$\text{where} \quad w_n = \frac{\frac{1}{\sigma_n^2}}{\sum_{m=1}^{N} \frac{1}{\sigma_m^2}} \tag{4.4}$$

and D_n are the observed magnitudes, with uncertainty σ_n and w_n is the normalised weight of each datapoint. We then make the parametrisation:

$$H = \frac{-Lq}{1-q} = -\frac{s}{1-r} \quad L = \frac{s}{r}$$

where $s = \sum_{in} w_n D_n$, $r = \sum_{in} w_n$ and q is the fractional transit length (assumed to be small), which gives:

$$\chi^2 = \sum_{all} w_n D_n^2 + \frac{s^2}{(1-r)^2} \sum_{out} w_n + \frac{2s}{1-r} \sum_{out} w_n D_n + \frac{s^2}{r^2} r - 2\frac{s}{r} s \tag{4.5}$$

By requiring that $\overline{D}$ is zero and as $\sum_{all} w_n = 1$ by definition we can make further substitutions for the out of transit summations:

$$\sum_{all} w_n = 1 \rightarrow \sum_{out} w_n = 1 - \sum_{in} w_n = 1 - r$$
$$\sum_{all} w_n D_n = 0 \rightarrow \sum_{out} w_n D_n = -\sum_{in} w_n D_n = -s$$

and find that

$$\chi^2 = \sum_{n=1}^{N} w_n D_n^2 - \frac{s^2}{r(1-r)} \tag{4.6}$$

As the first term is a constant and with the definition of χ^2 we can use the second term to construct a test statistic:

$$T = \left[\frac{s^2(in)}{r(in)[1-r(in)]}\right]^{\frac{1}{2}} \tag{4.7}$$

The test statistic is calculated for each set of in-transit points as defined by the test parameters. The maximum T gives the best parameter set. An estimate of the depth of the transit, $d \equiv H - L$, can also be recovered from the test statistic by parametrizing d we find:

$$d = H - L \rightarrow d = \frac{-s}{r(1-r)}$$
$$T = \left[\frac{s^2}{r(1-r)}\right]^{\frac{1}{2}} \rightarrow T = d\sqrt{r(1-r)}$$

The box-fitting method assumes that the periodic signal is a sharp box shape.

4.1.2 Matched filter

The matched filter algorithm relies on the comparison of a known signal to white noise. This requires a calculation of the probability that the observed signal is due to white noise (σ) or white noise plus a signal (S_n). The derivation of the test statistic

(Equation 4.8) from this ratio can be found in Kay (1998). The test statistic is:

$$T = \left(\sum_{n=1}^{N} \frac{D_n S_n}{\sigma_n^2} \right) \tag{4.8}$$

where D are the observed magnitudes, with uncertainty σ and S the test signal. To search for any signal in the dataset, S must be modified across the entire searchable parameter space. The highest resultant T is computed from the best set of parameters.

4.1.3 Deeg's method

In a attempt to detect a transiting extrasolar planet around the eclipsing binary CM Draconis the matched filter algorithm discussed in §4.1.2 was heavily modified to compensate for the unusual dataset. As the total dataset used in the research comprised several years of observations on several telescopes immediate comparisons between such could not be made. This method, described in Doyle et al. (2000), adds to the matched filter method a time-based weighting and while designed specifically for the CM Draconis dataset is applicable to any similar dataset.

Working with a single night's data at a time, the modelled stellar eclipses are subtracted from the lightcurve. The resulting lightcurve, which still contains extinction variations, has a parabola fitted, f^e, and this is compared to the original (eclipse subtracted) data giving the residual $r^e = |D - f^e|$. Each test signal is then subtracted from the lightcurve, and the residual is then fitted with another parabola, f^p. The fit of this parabola is then also compared to the original data to obtain the residual $r^p = |D - S - f^p|$. Comparison of the residuals indicates whether the with planet case (f^p) is a better fit to the data than the without planet case (f^e). Using

the two residuals, r^e and r^p a test statistic is calculated:

$$T = \sum_{n=1}^{N} \kappa_n \tag{4.9}$$

$$\begin{aligned} \text{where } \kappa_n = 0 \text{ if } \mathrm{t_{n+1}} - \mathrm{t_n} > 10\mathrm{min} \\ \text{and } \kappa_n = \frac{(r_n^e - r_n^p)}{\Delta t \sigma_D} \text{ otherwise} \\ \text{with } \Delta t = (t_{n+1} - t_n) \end{aligned} \tag{4.10}$$

where σ_D is the noise level for the night analysed. Each element of the test statistic is weighted by a time factor, this neglects points where the time spacing is too long and places more emphasis on those points which are close together. In the CM Draconis dataset with observations from multiple telescopes it accounts for the differences in time spacing.

4.1.4 Bayesian method

Unlike the previous methods the approach here results in a reconstruction of the transit shape. The periodic signal is represented by a Fourier series and the Fourier coefficients that best represent the signal are then found. The first step is to determine the most likely period of the signal. The global likelihood function (L) as derived in Defaÿ et al. (2001) from Baye's theorem:

$$\log L(\omega) \propto \frac{-N}{2\sigma^2} \sum_{k=1}^{m} \left(\frac{\alpha_k^2}{N^2} + \frac{\beta_k^2}{N^2} \right) \tag{4.11}$$

$$\begin{aligned} \text{where } \alpha_k = \sum_{n=1}^{N} D_n \cos(k t_n \omega) \quad k = 1...m \\ \text{and } \beta_k = \sum_{n=1}^{N} D_n \sin(k t_n \omega) \quad k = 1...m \end{aligned} \tag{4.12}$$

is maximised to find the best frequency of the signal. The signal shape can then be reconstructed from the Fourier coefficients:

$$f(t) = \sum_{k}^{m} = 1 \left(\frac{2\alpha_k}{N} \cos(kt\omega) + \frac{2\beta_k}{N} \sin(kt\omega) \right) \tag{4.13}$$

While theoretically an infinite Fourier series is required to fit the signal exactly, in Defaÿ et al. (2001) the authors indicate that the summation can be truncated at $m = 7$ without loss of precision. The method described here will only work for evenly spaced data - there is a strong dependence on the observation time, and the system was designed for closely spaced data.

4.1.5 Comparison

The data used in this research is spaced unevenly and not at all finely. The Bayesian method is therefore not at all suitable for this data. Deeg's approach on the other hand is quite suitable for unevenly spaced data however it also relies on nightly lightcurves which makes it unsuitable for the pilot dataset. Both the BLS and matched filter would be suitable for this dataset.

4.2 Maximised matched filter

The test statistic chosen for use in this research is the maximised matched filter, Tingley (2003b). This is a modified version of the matched filter detailed in §4.1, Equation (4.8). The matched filter test statistic is modified here to return a term dependent on the number of in-transit points, see Kay (1998) for the full derivation. In the case of a transit search we trial a number of different transit signals. To compare test statistics from different test signals, dependence on the number of in-transit points is essential. When this term is included in the derivation of T we

find:

$$T = d\left(\sum_{in}\frac{D_n}{\sigma_n^2}\right) - \frac{1}{2}d^2\left(\sum_{in}\frac{1}{\sigma_n^2}\right) \tag{4.14}$$

This equation then contains depth as a free parameter. This adds unnecessary complications to the calculation and thus to remove depth the equation is maximised with respect to d:

$$\frac{dT}{d(d)} = \left(\sum_{in}\frac{D_n}{\sigma_n^2}\right) - d\left(\sum_{in}\frac{1}{\sigma_n^2}\right) = 0 \tag{4.15}$$

Substituting for d into Equation 4.14 gives the resulting test statistic:

$$T = \frac{1}{2}\frac{\left(\sum_{in}\frac{D_n}{\sigma_n^2}\right)^2}{\left(\sum_{in}\frac{1}{\sigma_n^2}\right)} \tag{4.16}$$

In a transit search the test statistic T is calculated for each test signal. This is accomplished by determining which points are 'in-transit' for each signal and then calculating T. The maximum T for each lightcurve is determined. This method does rely on the assumption that out of transit points have an average of zero.

4.2.1 Advantages and disadvantages

The maximised matched filter (MMF) was the search algorithm chosen for this research. It creates a test statistic from three parameters, orbital period, transit length and transit phase. The important thing to note is that the test statistic contains a term which depends on the number of in-transit points. When comparing the likelihood of events that may have been observed a varying number of times - important in the case of unevenly sampled data, this term is necessary. The box-fitting technique (BLS) is very similar mathematically to the matched filter but includes values of points both in and out of the transit. In our case we are working with data points of δflux/flux where the average of the data points (which

is most of the points even with a transit) is zero. The elimination of the out of transit points speeds the computation by shortening the summation which is useful in the case (as it is here) where there is an extremely large dataset. The main disadvantage of the MMF, and indeed this holds for BLS too, is the squaring of the in-transit observations. This eliminates the difference between brightening and dimming events in the lightcurve, causing a larger occurrence of false positives. It is possible to remove this problem by testing whether the sum of the un-squared in-transit points is positive (dimmer) or negative (brighter).

4.2.2 Implementation

The test statistic is the basic mathematics of the algorithm and requires input of all possible parameter sets to then find the maximum as required by the statistic. Due to the maximised matched filter test statistics requirement of in-transit points for the summation, a separate piece of code is used to determine which points will fall within the transit time for a given parameter set. These data points are then fed into the final piece of code wherein the test statistic is calculated for that parameter set.

The external script is the code which determines how the stepping through the parameter space in all three variables is accomplished and also the way in which the maximum test statistic is determined. Different methods can be used for this.

4.3 Transit search maximisation methods

The transit algorithm is used to process the data and then the maximum test statistic is found. At first a simple fine grid search over period, starting phase and fractional transit length was trialled to find the maximum value of the test statistic, T_{max}. However, through testing using a small number of inserted artificial transits, this was found to be very computationally intensive and consequently other maximisation

methods were investigated.

4.3.1 Differential evolution

The method of differential evolution (Storn and Price, 1997; Price et al., 2005) was trialled to find the maximum test statistic using the 2002 data. Differential evolution (DE) is a genetic algorithm that searches for a global minimum in a multidimensional space. The algorithm used in this MATLAB code[1] minimises an externally defined objective function (i.e our test statistic). As the test statistic for the MMF algorithm is maximised here, we define the test statistic to be negative to ensure that the minimization matches with the maximisation required by the algorithm.

DE is used in conjunction with the transit detection algorithm to compute the best parameters for a transit in the data. The DE algorithm utilises random sets of starting parameters, sends these to the transit detection algorithm which returns T. Instead of stepping through the parameter space the algorithm then adds the weighted difference vector between two parameter vectors to a third new vector. This new vector has the test statistic calculated and which is then compared to the previous maximum T. If the T value is larger than the previous then the new vector replaces the comparison vector and then another new vector is found.

In this method the parameters are forced towards the result which gives the maximum value of T. It is possible for the DE algorithm to end up with a local maximum rather than the global maximum. The set up of the DE code used here allows the maximum and minimum values of the parameter set to be defined. In the case of the 2002 lightcurves, the parameters were set with a maximum period of 10 days, phase up to the total period length and maximum transit length of 0.25 days.

The differential evolution method was used on the pilot data (discussed in §5.2) but failed to achieve the expected results as the uneven spacing in the data

[1]Downloaded from *http://www.icsi.berkeley.edu/~storn/code.html* where it is freely available in MATLAB and other languages

caused the algorithm to find local maximum rather than the global maximum as sought.

4.3.2 Grid search

As discussed in §4.1 and §4.3.1 the test statistic, T (Equation (4.16)), is encased in a script to find the maximum value for T when the input parameters are modified. In the case of a grid search the script is the simplest possible. For each lightcurve the maximum test statistic is calculated through nested loops of period, phase and transit length. The in-transit points are determined and the test statistic calculated and then compared to the previous maximum value (or initial zero value). Providing the new value for T is larger than the previous maximum it is recorded with the parameter set that created it and then the loops continue.

The grid search is computationally slow due to the nested loops approach. The test statistic value must be calculated for every different parameter set, compared and then either kept or discarded. The number of different trial parameter sets determines the overall speed. In order to obtain the best results, the step size for each parameter needs to be small, however this requires a large amount of computation time and as such is not practical. A balance needs to be struck between the parameter step size and the amount of computation time. With limited time available for computations the grid search can be made coarser to minimise computation time, and any possible transits can then be re-analysed with a finer grid to obtain the best possible parameter set. However, such an approach is liable to miss transits.

4.3.3 Sparse matrix

In order to eliminate problems with speed while retaining a fine step size, the transit search was eventually implemented by means of a sparse matrix. Rather than using a set of nested loops to cover the parameter space at some particular phase resolution, a matrix formalism has been used to test every possible set of in-transit points

for a given trial period. The use of sparse matrices in MATLAB allows a considerable improvement in speed over the traditional grid search. A selection matrix S (Equation (4.18)) is formed where each row consists of a series of 1's and 0's where 1's mark a set of trial in-transit points. All possible sets of in-transit points, up to a given maximum number, are contained in S. The differing start point for each line gives the initial phase. If we define $y_n = \frac{D_n}{\sigma_n^2}$ and $w_n = \frac{1}{\sigma_n^2}$ then Equation (4.16) becomes

$$t_i = \frac{1}{2}\frac{(S \cdot \underline{\mathrm{y}})_i^2}{(S \cdot \underline{\mathrm{w}})_i} \tag{4.17}$$

where $\underline{\mathrm{t}}$ is a vector containing the test statistic corresponding to each row of S. Thus from a single set of matrix operations we determine the best transit length and initial phase for a particular period and loop only over periods. This cuts down on processing as the code contains a single loop rather than three, and the sparse matrix arithmetic is implemented extremely efficiently internally in MATLAB.

$$S = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix} \tag{4.18}$$

4.4 Variable star detection

Variable stars are generally easier to detect than transiting extrasolar planets. The amplitude of variations in a variable star is often substantially larger than the depth of a transit. Variable star amplitudes range from thousandths of a magnitude to 20 magnitudes (novae). In most cases the lightcurve will be continuously varying with phase. A transit, by contrast, varies from constant light for only a small fraction of the lightcurve.

There are a number of period detection algorithms from all disciplines that can be used for detecting and determining periods of variable stars. Most immediately the difference images can be examined for changes in brightness - this is suitable if one is only examining a small number of stars and unless the star is extremely well sampled a period determination algorithm is still required. If there is a single star which is suspected to vary and an evenly spaced set of observations of this star, almost any period determination algorithm can be used.

For data such as involved in this research an algorithm suitable for use with unevenly spaced data and convertible for use with a large dataset is needed. The most well-used algorithm for unevenly spaced data is the Lomb-Scargle algorithm §4.4.2. It is also possible to use the Lomb-Scargle algorithm in a similar way to the transit detection algorithm, with a test statistic, in this case the false alarm probability, which can be given a threshold below which periods are most likely to indicate true variability. The Lomb-Scargle algorithm is however designed for sinusoidal variations and may have trouble with the variations of eclipsing binaries. For non-sinusoidal variables, and possibly also transits, the Phase Dispersion Minimization algorithm is suitable, see §4.4.3. Variable star detection can then proceed in a similar manner to transit detection.

4.4.1 Image detection

The simplest variable star detection method is examination of the difference images. A variable star will not contain the same amount of light as it has in the reference frame through the whole dataset, it will not therefore be completely eliminated when subtracted by the reference frame. The difference images will be left with residual flux from the variable stars. In a single difference image many of the variable stars are too faint to be readily discernable. In order to overcome this a number of images are summed together. The subtracted images are squared and then summed to create a variance image. Taking a square-root of this produces a standard deviation image. The non-varying stars will have minimal residuals post-subtraction and in the standard deviation image will not show as very different from the background. A variable star by contrast will show up as a bright object and will be easy to detect. Looking at the standard deviation image, Figure 4.1 it is easy to discern the variable candidates. Image detection cuts down on the number of stars which are processed by a period search algorithm since only those detected from the summed image are searched.

Problems arise with the saturated stars, which leave large residuals on the subtracted images. Summing the saturated stars causes large clumps of pixels which are not star shaped but have a large enough photon count to be detected by automated methods. Two ways around this are proposed: to do the variable detection from the images by hand; to find a method to mask the saturated residuals. By eye detection may miss faint or low-amplitude variables and has a large dependence on the display settings. Masking the actual saturated stars is relatively easy but masking the external residuals is not. A secondary way to lessen both these problems is to sum only those images with good seeing. The external residuals around the saturated stars will be lessened and there will be less noise in the variations allowing the variables to stand out more from the background.

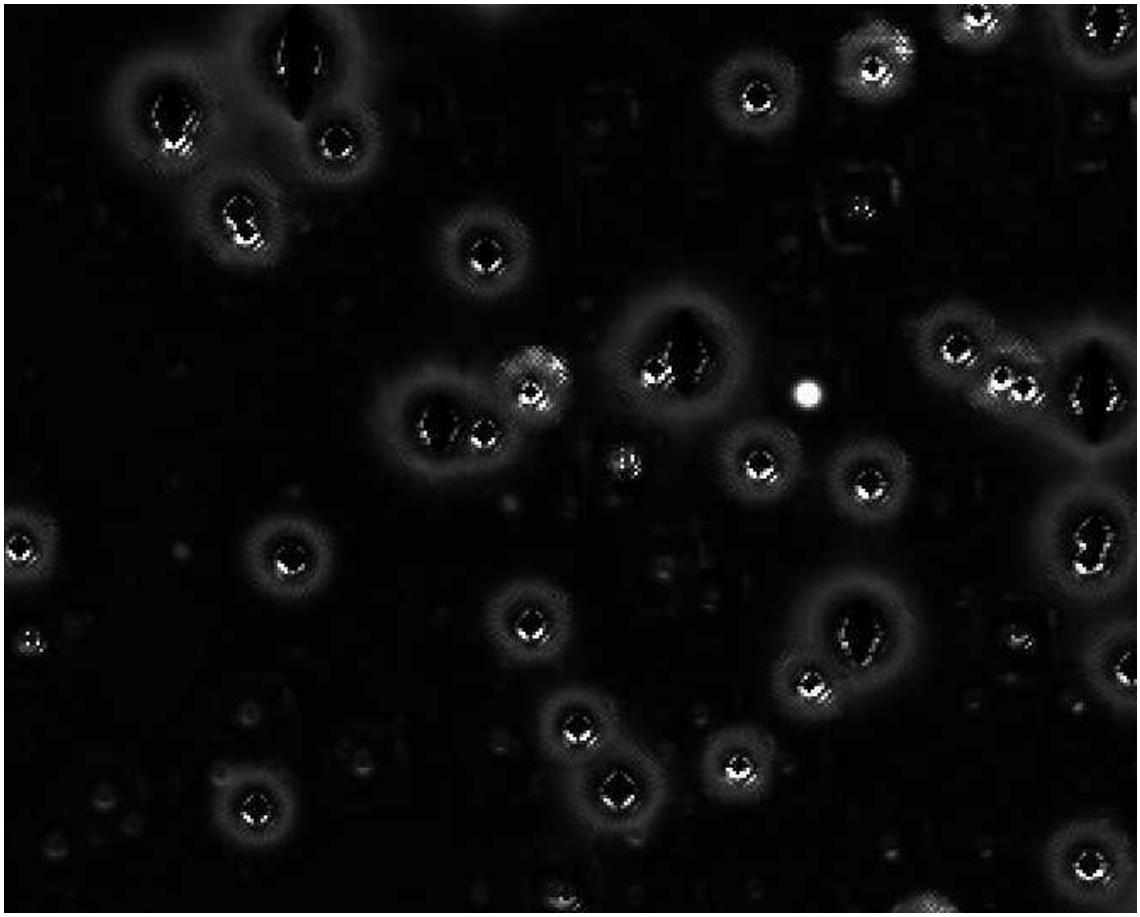


Figure 4.1: **Variable star detection by image:** Detection of a variable star can be performed on summed difference images. In this frame a single variable star is visible as a white circle while annular residuals from the subtraction are all that can be detected of other constant stars.

Detecting variable stars from the images is certainly doable. A test run detected around 15 variables from an image containing 10,000 stars. However these stars and more were also easily determined from the search algorithms detailed next. The image detection method requires a lot of personal input and thus can be inefficient when compared to computational methods.

4.4.2 Lomb-Scargle algorithm

The Lomb-Scargle algorithm (henceforth LSA) is based on the fast Fourier transform (FFT) period determination approach. A series of sinusoidal waves are tested as best fit to the data. FFT has limitations when it comes to unevenly spaced data such as the dataset used in this research. The LSA was developed specifically to combat this, (Scargle, 1982). A frequency or period range is submitted to the search algorithm. This is converted to angular frequency (ω) and a power spectrum calculated for the lightcurve. The power as a function of frequency (ω) is defined by Equation 4.19 and an example of the output is seen in Figure 4.2.

$$P_X(\omega) = \frac{1}{2}\{\frac{[\sum_{j=1}^{N_0} X(t_j) cos\omega(t_j-\tau)]^2}{\sum_{j=1}^{N_0} cos^2\omega(t_j-\tau)} + \frac{[\sum_{j=1}^{N_0} X(t_j) sin\omega(t_j-\tau)]^2}{\sum_{j=1}^{N_0} sin^2\omega(t_j-\tau)} \qquad (4.19)$$

$$\text{where } \tau \text{ is defined by} \qquad tan(2\omega\tau) = \frac{\sum_{j=1}^{N_0} sin2\omega t_j}{\sum_{j=1}^{N_0} cos2\omega t_j}$$

$$S_X = 1 - (1 - e^{-P_X})^N \qquad N = \text{ number of frequencies} \qquad (4.20)$$

The frequency giving the highest power value determines the most likely period of a variation and a false alarm probability value is calculated which gives an indication of how likely this value is a true periodic variation (S_X, Equation 4.20).

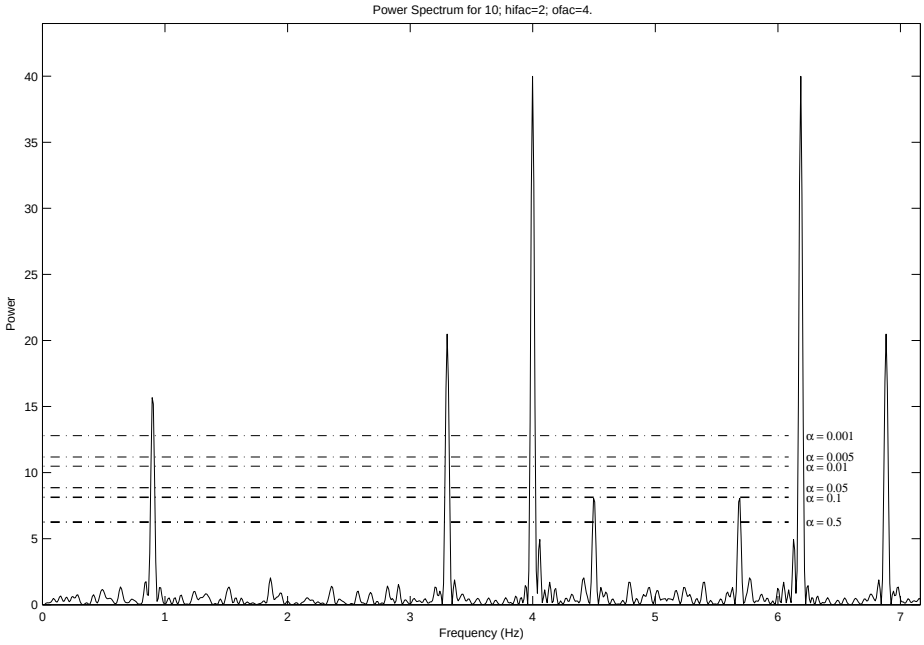


Figure 4.2: **Lomb Scargle Power Spectrum** Output from the Lomb-Scargle algorithm with a simulated lightcurve. Lowest false alarm probability (α indicates best possible fit.

4.4.3 Phase dispersion minimisation

In an article published by Lafler and Kinman (1965) a technique of period determination was proposed to deal with a survey of RR Lyrae stars. This method was designed to search for periods by use of an electronic computer (at this stage a rather new idea). Simply the technique was an automated version of the standard method - the period producing the least observational scatter around the mean is classified as the best fit period. This technique is well suited to cases in which a non-sinusoidal time variation is observed irregularly for a short period of time, this makes it a promising algorithm for searching our data for both variables (a large number of which are expected to be eclipsing binaries) and transits. Stellingwerf (1978) developed a generalization of the 1965 technique, which greatly increased the ease of use of this method.

Our dataset has a set number of observations observed at varying cadence in two blocks over three years. The search for periodic signals in this data fits the criteria in which the search method of Stellingwerf is useful.

Given a set of observations represented by the magnitudes, D, and the times, t, the variance is defined by Equation (4.21), where $\overline{D}$ is the mean; $\overline{D} = \sum D_i/N$.

$$\sigma^2 = \frac{(\sum(D_i - \overline{D})^2)}{N-1} \tag{4.21}$$

If we then choose M distinct samples of datasets (containing n_j datapoints) each with variances s_j^2 $(j = 1, M)$, the overall variance for all the samples is given by

$$s^2 = \frac{(\sum(n_j - 1)s_j^2)}{\sum(n_j - M)} \tag{4.22}$$

For a particular trial period P we compute the phase vector $\Phi : \phi_i = t_i/P - int(t_i/P)$ and then choose M samples from the lightcurve with the criteria that sample j have members all with similar ϕ_i. This is most easily completed by binning the full phase

but can be completed in any way such that the criterion are satisfied. The statistic is then given by

$$\Theta = \frac{s^2}{\sigma^2} \tag{4.23}$$

If P is found to be a correct period, Θ will reach a local minimum as compared to neighbouring periods. For implementation a series of periods are trialled and the lowest dispersion as determined by calculating Θ gives the most likely period.

In our implementation the variation seen in Equation (4.24) is used.

$$\Theta = \frac{\sum_j s_j^2}{\sigma^2} \tag{4.24}$$

which eliminates the need to calculate overall variance of the samples.

The search algorithm calculates the statistic Θ (Equation (4.24)) for each lightcurve by taking running samples of 3 consecutive phase points, created by dividing the time by the current period, each having variance s_j and minimising the test statistic, Equation (4.24), where σ^2 is the overall variance of the lightcurve. This produces a best period for the lightcurve. If we apply the algorithm to a transit search, a χ^2 fit is then used to determine the best phase and transit length for the best period.

Chapter 5

Transit candidates

As described in Chapter 1, a lightcurve of a transiting extrasolar planet is characterized by a small dip in the amount of light reaching the observer from the host star. The transit detection algorithm is applied to the dataset and the results analysed for lightcurves which have transit-like properties.

5.1 Testing the transit search algorithms

In order to understand how the search algorithms work and the detection thresholds for varying parameters the algorithms are tested. Simulated transits are the simplest manner in which to accomplish this. Before the real dataset was available artificial lightcurves were created with inserted transits and searched by the algorithms. In order to test lightcurves with realistic noise properties, and thus make a more testing trial, a secondary (previously searched) dataset was acquired. Trial runs of the algorithm were performed both on completely artificial lightcurves - this tested the coding, and then on the secondary dataset, §5.1.1. When our complete dataset was available, artificial transits were inserted into a subset of this to determine suitable threshold values for the detection algorithms, §5.1.2.

5.1.1 Artificial transits in other data

To test the grid search algorithm a number of artificial transits of different periods and depths were inserted into a set of relatively constant lightcurves from the pilot dataset. These modified lightcurves were passed into the algorithm with the original lightcurves and others from the data set. The change in position of the test statistics of the modified lightcurves with respect to the main body of results was large, Figure 5.1. This initial testing allows the determination of where to start looking for lightcurves containing transits.

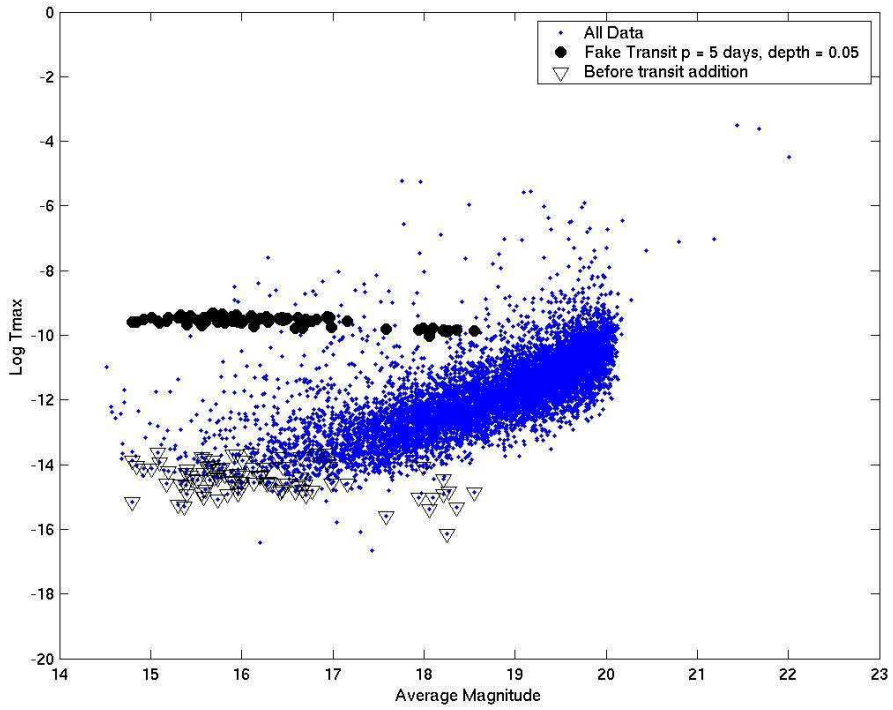


Figure 5.1: **Artificial transits** The results from insertion of an artificial transit into a selection of lightcurves from the pilot dataset. Each inserted transit had a period of 5 days, depth of 0.05 and a transit length of 5 hours. The change in position of the test statistic value allows for a better understanding of the lightcurves to follow-up for transits.

In July 1999 observations were made by the Hubble Space Telescope (HST) of

the globular cluster 47 Tucanae (Gilliland et al., 2000)[1]. These observations resulted in a large amount of high precision photometric data suitable for planetary transit detection. This data is perfect for testing transit detection algorithms. In particular we have used it to test the detection efficiency of different methods and to examine the effect different parameters have on the test statistic.

5.1.2 Simulations of transits in the full dataset

To determine an appropriate threshold, simulated transits were placed within a subset of the full dataset. Random lightcurves were chosen to have transits inserted. Transits were simulated at differing periods and depths. Random periods were chosen between 1 and 5 days with a variety of depths from 1% to 6%. The initial starting phase was also randomized. The transit length is varied in relation to the period using an approximation from the equations of Seager and Mallén-Ornelas (2003).

There are problems with the uneven spacing of the data that cause a large number of the inserted transits to have minimal effect on the lightcurve. Those with correctly detected transits have an obvious change in the test statistic. Figure 5.2 shows the data-points with correctly found transits before and after the addition of the simulated transits. The test statistics for the lightcurves without simulated transits are far lower than that for the lightcurves with simulated transits. All the test statistics for lightcurves without simulated transits fall below the threshold value.

[1]Thanks to Ron Gilliland for supplying the lightcurves

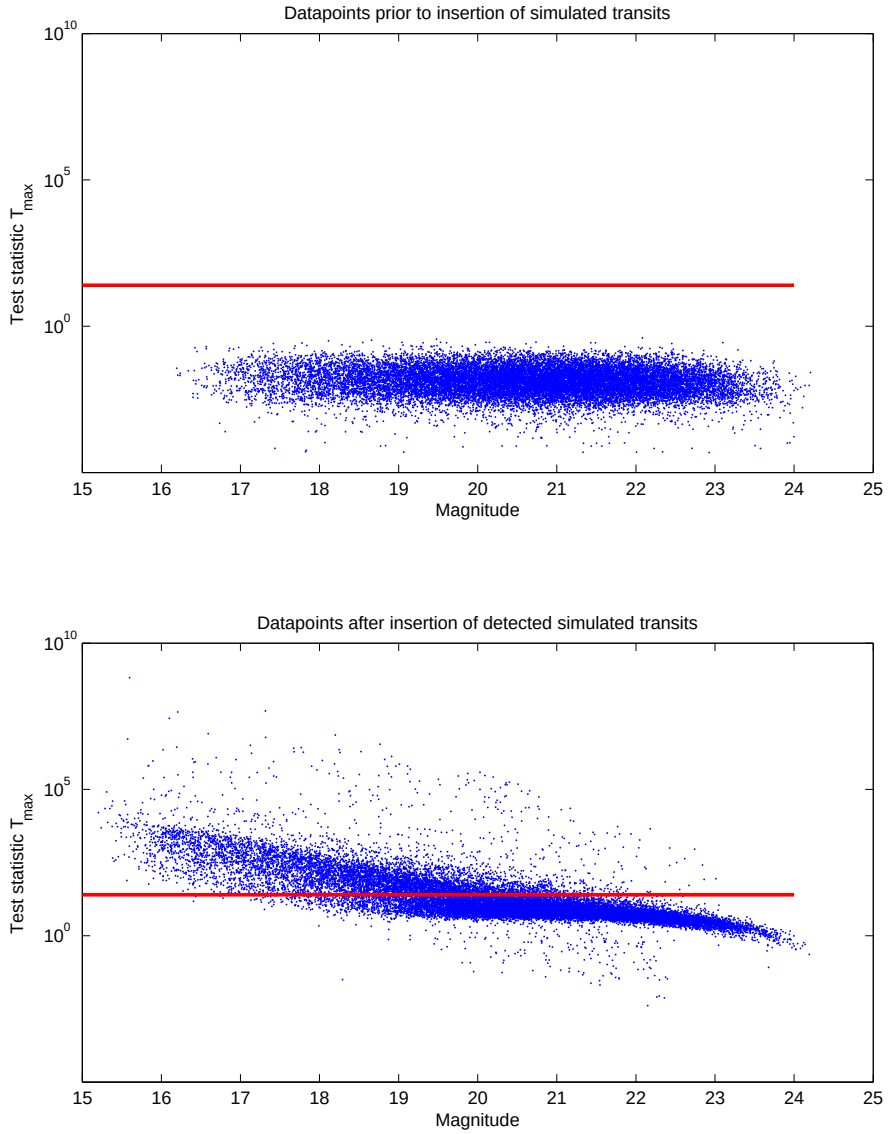


Figure 5.2: **Results of transit simulations** Results from insertion of simulated transits into real data and searched with the transit detection algorithm. The top panel shows the test statistics of the lightcurves prior to transit insertion. The bottom panel shows the same lightcurves with test statistics calculated after the insertion of a simulated transit. The bold line shows a threshold value of 25 for reference.

From the results of the simulated variables we are able to determine detection efficiency (§5.3.1) and threshold values (§5.3.2) for the transit search.

5.2 Initial transit search

An initial transit search was performed on the pilot dataset. This was obtained in 2002 and consists of approximately 350,000 lightcurves, each with 63 data-points. Analysis of the pilot dataset proved unsuccessful in finding transits. This is not surprising given the small number of data points. From the transit simulations discussed previously we know that a detected transit should have a test statistic above the threshold, that is above the main body of results. From Figure 5.4(a) it we see that a number of the lightcurves fall into this area. Some of these lightcurves belong to variable stars (e.g. Figure 5.3). These are clearly not what we are searching for however a number of these appear in the results. Also in the results there are a number of lightcurves that have large variations in intensity due to variations in image quality rather than within the stars. In order to make the best use of the data (and before the decision was made to make new observations) manipulations of the pilot dataset were made in an attempt to make detections.

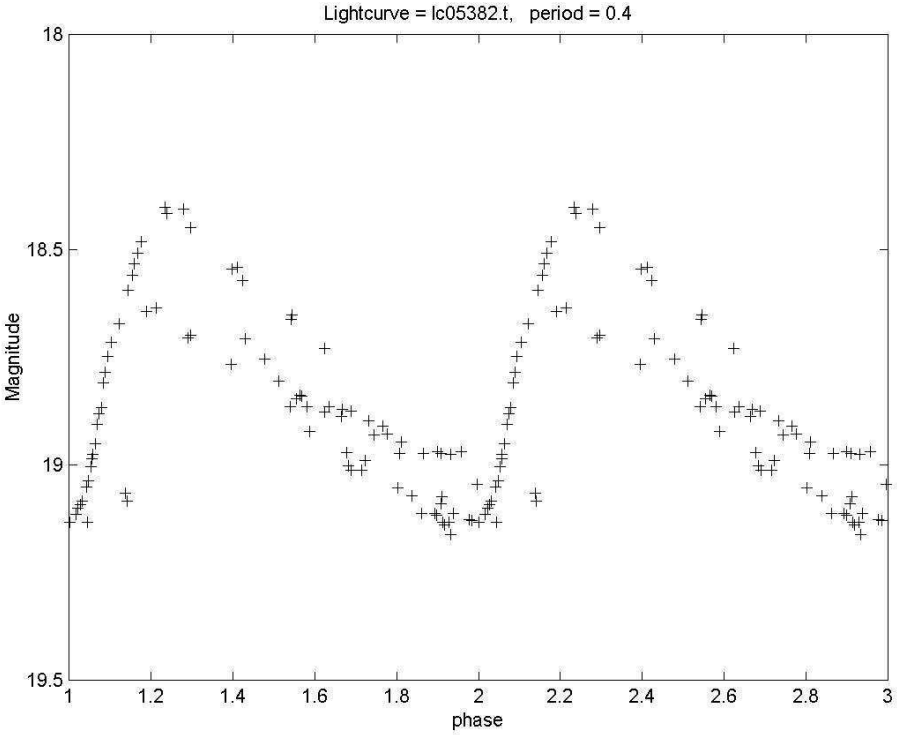
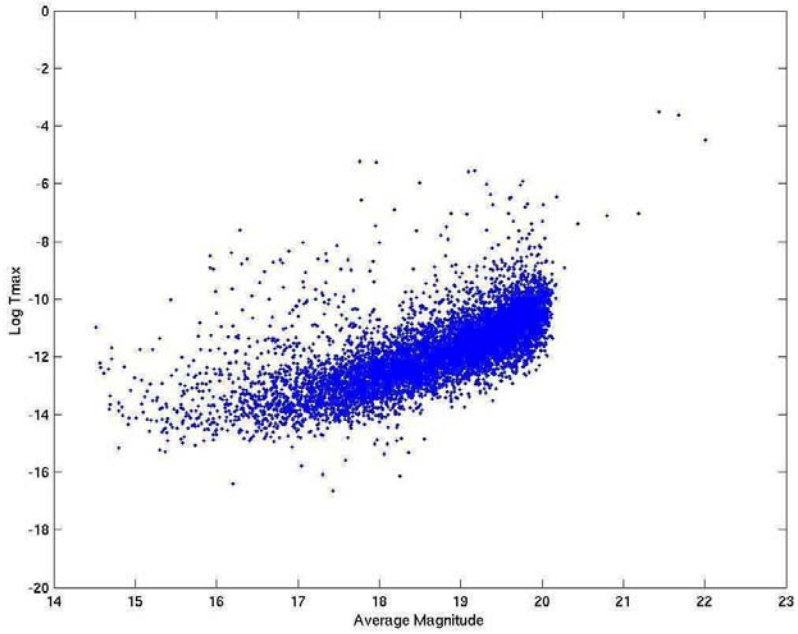


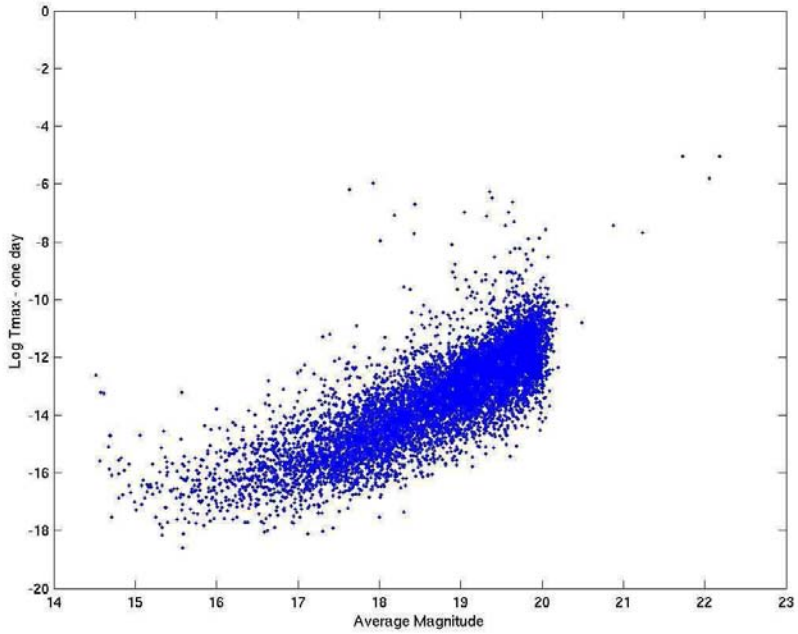
Figure 5.3: **Detection of a variable star** Lightcurve of a variable star found in the pilot dataset using the transit detection algorithm with differential evolution maximisation.

5.2.1 Peak time data

The pilot dataset was collected in the manner of a microlensing event, the data is spread across a month but the majority of the data points are found within two days. The results from processing these heavily sampled days are shown in Figure 5.4(b). The results are very similar to that found from the total dataset, Figure 5.4(a). This subset of data has a lesser standard deviation than the total dataset and a lightcurve containing a transit should stand out more due to the lowered standard deviation of those points in the main section. This appears the case - there are fewer points found above the main body.



(a)



(b)

Figure 5.4: **Test statistic from detection algorithm** The top panel shows the test statistic determined from using the differential evolution algorithm on a subset of the 2002 data. The bottom panel shows the same subset of 2002 data but the search algorithm only used those data points collected in two heavily observed days.

The differential evolution method (§4.3.1) even on the one day of data did not have a high detection efficiency even on very deep artificial transits. Running this search on the pilot dataset produced no transit candidates. This is most likely due to the small number of data points in each lightcurve and the uneven time spacing. At this point the pilot dataset was determined to be unsuitable for transit searches and a secondary dataset of the same field was acquired, §2.3. At the same time another implementation of the transit search, the sparse matrix, was made, §4.3.3.

5.3 Implemented search algorithm

With the problematic results from the differential evolution method and the pilot dataset, the total dataset was searched using the sparse matrix method discussed in §4.3.3. This trials every possible set of in-transit points for each period. The sparse matrix search was run on the de-trended dataset, §3.8.2 and the un-de-trended dataset. Table 5.1 shows a sample of the results matrix from the search. Figure 5.6 shows the test statistic results from running the transit search algorithm on different subsets of lightcurves. Each subset corresponds to a different 1k by 1k sub-image region of the field.

Lightcurve	RMS	Instrumental Magnitude (R)	Resultant Period	Resultant Phase	Transit Length	Test Statistic
6400	NaN	NaN	0	0	0	0
6401	0.206324	17.024	1.9305	0.767366	0.0048174	994.436
6402	0.118437	16.225	2.2998	0.133794	0.00200017	4.23635
6403	NaN	NaN	0	0	0	0
6404	0.0824703	15.981	1.4233	0.527366	0.00583152	3.34305
6405	0.0300639	14.885	1.6881	0.255435	0.0363723	3.77377
6406	0.0118495	13.326	1.9854	0.490531	0.00962023	7.5468
6407	0.231454	16.678	2.3179	0.52828	0.0178179	9.10823
6408	0.0160204	13.785	1.9181	0.082686	0.0636567	7.99897
6409	0.0367927	14.692	1.2801	0.102101	0.0323412	9.64506
6410	0.0140042	13.534	1.9195	0.1677	0.0131284	6.00883
6411	0.00351449	10.874	0	0	0	0
6412	0.0474845	15.008	2.2532	0.601944	0.0557429	7.36575
6413	0.00311493	10.959	0	0	0	0
6414	0.228983	17.117	1.2035	0.10752	0.0351475	1.37108
6415	0.106712	9.993	0	0	0	0
6416	0.128238	16.491	3.1073	0.656132	0.0137418	4.03272
6417	0.0125254	12.266	0	0	0	0
6418	0.0798397	15.692	1.649	0.860582	0.00824742	5.82552
6419	0.311336	17.518	2.3611	0.656135	0.00266825	3.3739
6420	0.0284575	13.869	2.14	0.0530841	0.0253738	12.6588

Table 5.1: **Transit search results matrix** Sample of results as shown in the matrix resulting from the sparse matrix algorithm on a subset of lightcurves. Columns are star number, period, initial phase, transit length and test statistic.

5.3.1 Detection efficiency

From inserting approximately 10,000 artificial transits into subsets of the full dataset (2002 and 2005 data) we can determine the detection efficiency of the search algorithm. Due to the uneven spacing and limited number of data points in each lightcurve it is un-surprising that this is low.Figure 5.5 shows the recovery rate of inserted transits in real data. Each panel shows the fractional correct detections as a function of magnitude for a different inserted amplitude. We define a detection as being when the period found is within 10% of the inserted period or an alias and has a test statistic above the threshold (see §5.3.2 for more details on choosing the threshold). The detection efficiency peaks at around 50% in the brighter magnitudes. The detection efficiency then drops dramatically to the fainter magnitudes.

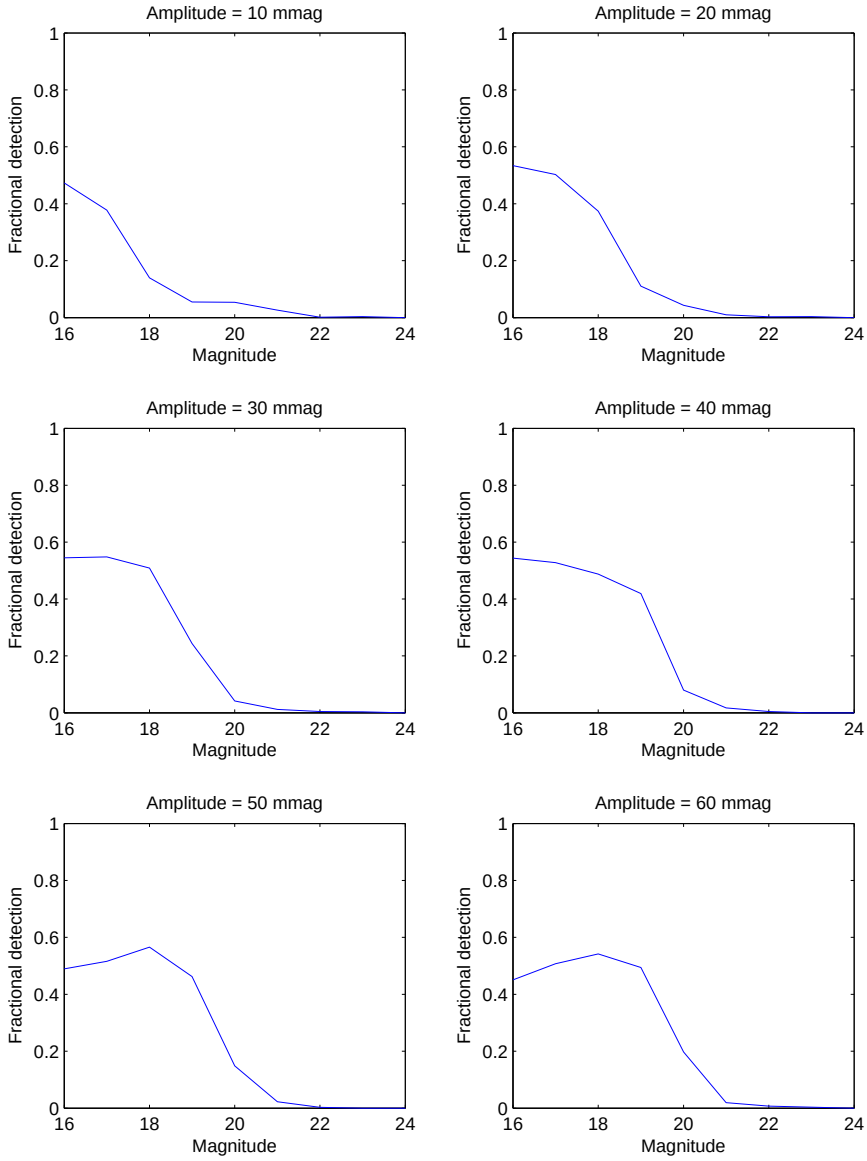


Figure 5.5: **Detection efficiency of transit simulations** Results from insertion of simulated transits into real data and searched with the transit detection algorithm - panels show correctly detected transits as a fraction of inserted transits as a function of magnitude for each amplitude.

5.3.2 Detection criteria

Due to the reduction process the dataset has been split into 64 subsets each containing approximately 8,000 lightcurves. In order to reduce the number of lightcurves examined by eye after the algorithm has been applied, threshold values were determined for each dataset. The threshold value for each subset is slightly different as it is based on the test statistic value being above the noise.

To determine a threshold value for each subset the test statistic for each lightcurve was plotted against the reference magnitude (see Figure 5.6 for examples). In the majority of cases the threshold was a simple cut at a value of T. Above this, the lightcurves were examined by eye (Figures 5.6(a,b)). In some of the data plotting the test statistic revealed faint stars with T values clearly above the majority but below a straight line threshold, (Figures 5.6(c,d)). In these cases a second sloped line was calculated and stars fainter than the turning point examined by this criteria instead.

Using these detection criteria allowed the examination of a much smaller number of lightcurves (approximately 3%). In a large number of cases examination by eye showed a transit-like signal, but deeper than an actual transit could be. However a significant number of lightcurves were found where there was a transit-like signal of appropriate shape and depth. As the algorithm is not dependent on depth this parameter is not calculated in the search. This can be calculated later where d = mean(out of transit) - mean(in-transit). This was not made part of the automated detection criteria but quick examination of the axes scale allowed the too deep candidates (depth more than 500 mmag) to be quickly rejected. A preliminary list included a large number of lightcurves with the deviation too deep (more than 200 mmag) or with only a single point deviation. These were eliminated on a secondary examination of the lightcurves leaving the final list as in §5.6.

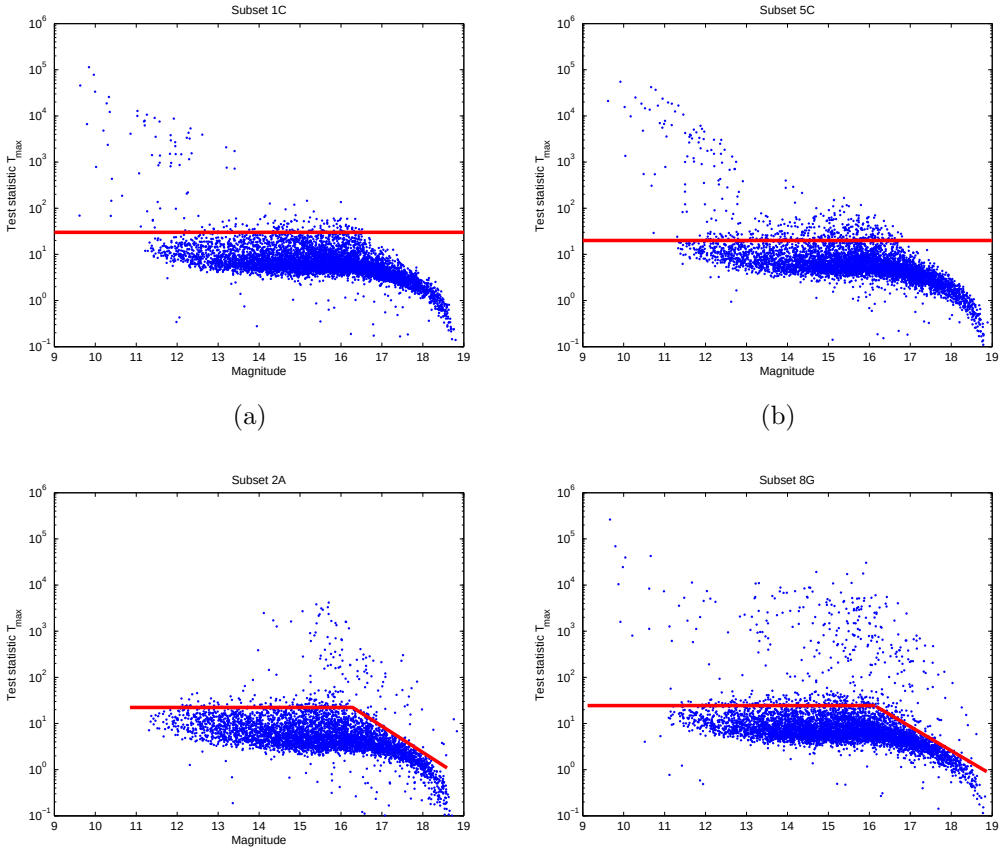


Figure 5.6: **Detection threshold of the transit search** Figures showing detection thresholds for several subsets of data. (a) and (b) show subsets where a single threshold was possible, (c) and (d) show subsets where multiple thresholds were used. Each figure is on the same scale, with the thresholds used represented by thick lines.

5.4 Transit search results

The transit search algorithm has resulted in a number of candidates. The initial pass included a large number of lightcurves with high test statistics resulting from non-transit features in the lightcurves or single point deviations. The secondary identification has resulted in a set of transit candidates numbering 31, some of which were then further eliminated as our definition of a candidate was refined. The initial transit candidate list only contained candidates found in the de-trended data as no good candidates were found in the un-de-trended data. The un-de-trended lightcurves did often show the same signal as the de-trended lightcurves when phased with the de-trended lightcurves period. However the search on un-de-trended data found other periods with higher test statistics.

After analysis of the lightcurves by the transit detection algorithm, those lightcurves with test statistics above the threshold value were examined by eye to test their veracity. From initial calculations the expectation was that there should be around 10 transiting extrasolar planets in the dataset (see §1.4.7.3). However our light curves contain only 267 (204 (2005) + 63 (2002)) data points compared to the 687 (624 (2005) + 63 (2002)) that we expected from our observing proposal. This resulted in a reduced detection efficiency and reduced expectations in the number of detected planets.

From the detection efficiency calculations we can make an estimate of how many planets we are likely to find. The initial prediction is for ten planets to be found across all magnitudes assuming a detection efficiency of 100%. Using the detection efficiency diagrams (Figure 5.5) we can determine the effective number of stars at each magnitude around which an existing planet could be detected. By summing over all magnitudes we determine a fraction of the total number of stars we are effectively searching and from there an estimate of the expected yield. Using the numbers from the 40 mmag amplitude (which were closest to an average) simulations

we find a effective number of stars of 89,000 or 16.59% of the total number of detected stars. The expected number of detected extrasolar planetary transits is therefore 1.66.

Transit searches however are likely to pull in more than this as candidates. Some of these will clearly not be extrasolar planets. It is likely that a number of the false positives will be blended eclipsing binaries or stars with undetected neighbours. In order to confirm transit detections each candidate must be carefully examined to determine its nature. The most accurate method is to follow-up the photometric observations with radial velocity measurements. These measurements will allow the size of the variation to be confirmed as likely to be a planet or a secondary star. The largest problem with such a method is the lack of suitable facilities to make measurements of the faint stars involved in this survey. There are other methods suggested for determining the nature of a transit candidate such as the colour of transit curve, which could be used with more photometric observations.

5.5 False positives and elimination methods

Transit detection surveys are plagued by false positives. It is entirely possible that every transit candidate selected from the algorithms is actually a false positive and not an extrasolar planet. There are a number of stellar systems which can produce lightcurves similar to that of transiting extrasolar planets.

The first method used to test the candidature was to examine closely the object in the images. This does not involve any extra observations and is easily accomplished. If it is an isolated star it is unlikely that there are problems with blends. If the star is extremely bright it is possible that there are undetected faint neighbours eclipsed by the primary star. If examination of the images provides no reasonable alternative explanation, further work is need to confirm the planetary nature of the lightcurve.

The first method of transit confirmation by means of further observations involves taking targeted photometric observations through the transit in multiple filters. The colour changes through a transit caused by an extrasolar planet are very distinct from that of a blended star or grazing eclipsing binary. The transit depth in the colour index (B - I) is around 15% that of the B or I depths alone. The depth of a blend in the colour index is far stronger than a transit signature. Proposed initially by Rosenblatt (1971) as a method to detect transits with less precise photometry than is available now, the two-colour method relies on a ratio between two colours to indicate the detection of a transit. Now, as proposed by Tingley and Sackett (2005), it can be used a method to confirm the presence of an extrasolar planet by measurements of the colour changes through a transit.

The other, and most ideal, method of confirmation of planetary transits is to make high precision radial velocity measurements of the star and confirm the planet in the same manner as radial velocity planets are detected. The combination of both measurements allows a large variety of parameters to be determined and the mass confirmed. If radial velocity measurements are taken as the planet transits the star measurements may also detect the Rossiter-McLaughlin effect. This allows measurements to be made of the difference in angles on the sky of the planetary orbital axis and the stellar rotation axis, giving in essence a measure of the inclination of the ecliptic and such measurements can be useful in tracking evolutionary paths. Problems with radial velocity follow-up occur with faint stars as there are not at this stage spectrometers available to measure the planetary Doppler effects ($\sim$ 5 ms^{-1}) on faint stars. However, medium-resolution spectroscopy can be used to set an upper limit on radial velocity amplitude and thus to reject many grazing eclipsing binaries.

5.6 Initial candidates

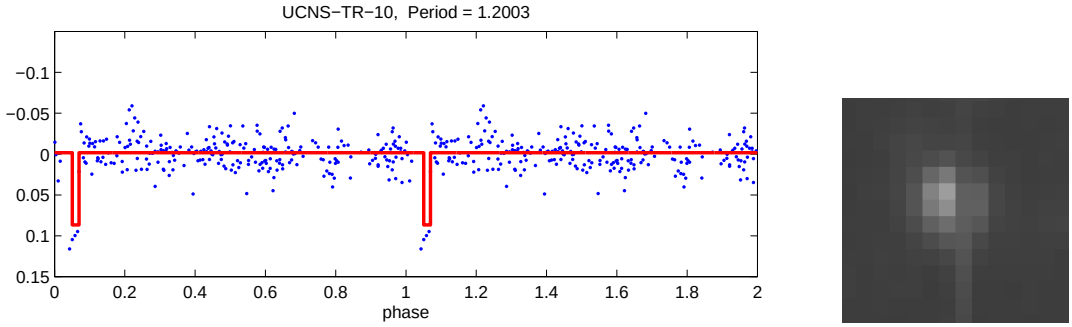
Following a preliminary search by eye of those lightcurves with test statistic above the threshold a number of candidates have been determined. The initial list contained 31 candidates which had lightcurves resembling those expected from extra-solar planetary transits. Further examination of these light curves removed eight of these from consideration. The final list of possible candidates contains 23. These candidates are listed in Table 5.2 and have been given identifications of UCNS-TR-# (University of Canterbury Norma Search Transit). The images of these candidates have been examined in detail to determine if there are any effects in the image that may have caused the transit-like lightcurve and those that pass this test were then re-reduced using code similar to that for the entire reduction but focussing on a single star as the code was originally intended when used in microlensing observations.

ID	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitude (R)	(B)	(I)	Period (days)	Image Notes
UCNS-TR-1	16:20:53.94	-53:22:37.45	17.77	20.45	15.49	1.4329	
UCNS-TR-2	16:20:52.33	-53:20:17.44	19.43	21.77	18.63	1.2840	
UCNS-TR-3	16:20:37.55	-53:21:00.08	18.67	20.94	17.86	1.4233	
UCNS-TR-4	16:20:53.27	-53:16:53.00	19.18	21.77	18.62	1.9224	neighbour
UCNS-TR-5	16:20:22.56	-53:16:24.37	18.53	21.76	16.95	1.7001	on edge MolCl
UCNS-TR-6	16:19:44.85	-53:22:52.05	19.13	20.97	17.94	1.7188	
UCNS-TR-7	16:19:46.61	-53:22:39.40	17.69	20.25	16.15	1.4121	
UCNS-TR-10	16:18:36.31	-53:24:02.30	19.25	21.82	18.15	1.2003	saturated overflow
UCNS-TR-11	16:18:26.78	-53:20:21.45	19.35	21.25	18.04	1.2016	saturated overflow
UCNS-TR-12	16:17:22.82	-53:26:03.25	18.85	21.58	17.95	3.7707	probable blend
UCNS-TR-13	16:18:02.98	-53:20:38.77	17.27	19.64	16.31	1.2016	
UCNS-TR-14	16:17:46.83	-53:21:35.21	17.90	19.98	17.08	1.2009	neighbour
UCNS-TR-15	16:17:18.28	-53:21:12.96	18.55	20.46	17.86	1.5504	
UCNS-TR-16	16:17:47.93	-53:14:51.30	17.62	18.96	17.23	1.3104	
UCNS-TR-17	16:17:37.37	-53:15:47.75	19.79	22.48	18.28	1.7274	
UCNS-TR-19	16:17:38.94	-53:37:57.06	19.35	21.64	18.99	1.7638	faint
UCNS-TR-20	16:17:43.49	-53:28:45.79	17.97	20.28	17.30	1.4234	
UCNS-TR-23	16:19:46.85	-53:39:15.40	17.24	18.97	16.76	2.7020	
UCNS-TR-24	16:19:20.24	-53:38:10.67	17.53	19.24	17.21	1.4259	
UCNS-TR-26	16:20:40.28	-53:40:07.11	18.25	20.14	14.14	1.5784	neighbour
UCNS-TR-28	16:17:23.50	-53:34:38.93	17.35	19.59	16.62	2.1354	neighbour
UCNS-TR-30	16:18:24.16	-53:29:02.38	17.93	20.07	17.45	1.2149	probable blend
UCNS-TR-31	16:20:17.14	-53:31:50.24	19.85	22.10	18.97	2.6269	neighbour

Table 5.2: **Initial transit candidates** Table of initial transit candidates prior to examination. Tabulated are identification, coordinates, R magnitude, period and notes from examining the reference image.

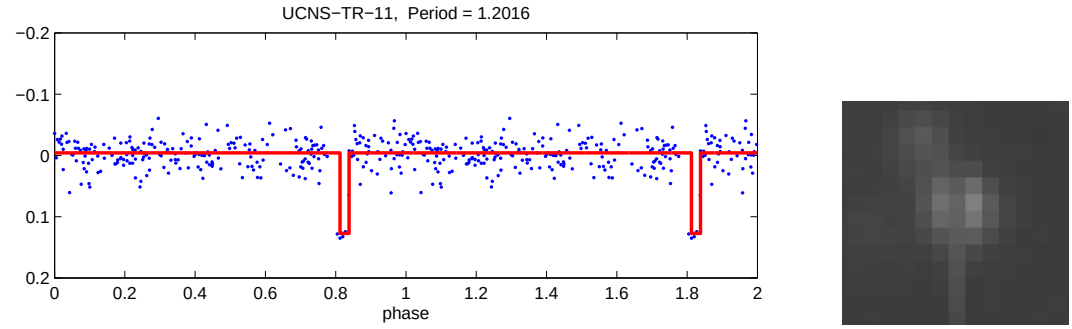
5.6.1 Rejections from images

Two of the transit candidates were eliminated from the list after examination of the reference frames. In two cases the candidate stars were contaminated in the reference frame by overflow from saturated stars. The lightcurves of these candidates (plotted without error bars) along with finder images are shown in Figures 5.7 and 5.8 with the in-transit difference images seen in Figure 5.9. Only some of the in-transit frames show the diffraction spike, the other frames show no signal above the noise, Figure 5.9.



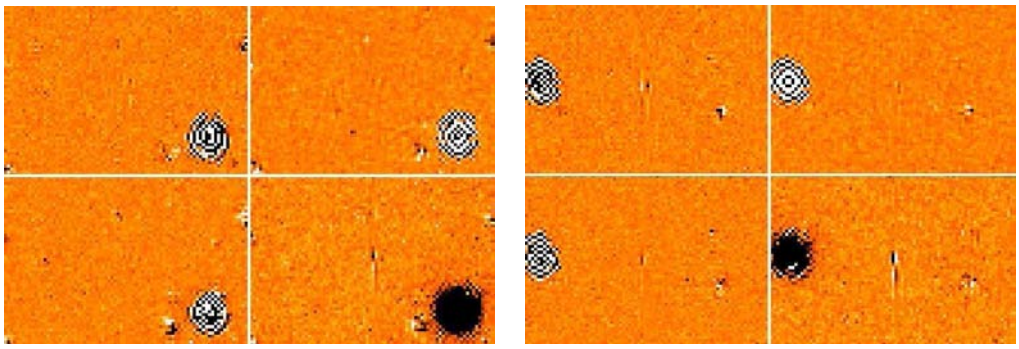
(a) (b)

Figure 5.7: **Rejected candidate UCNS-TR-10**: transit candidate lightcurve rejected due to image. Diffraction spike present in reference frame.



(a) (b)

Figure 5.8: **Rejected candidate UCNS-TR-11**: transit candidate lightcurve rejected due to image. Diffraction spike present in reference frame.



(a) (b)

Figure 5.9: **Candidate UCNS-TR-10 (a) and UCNS-TR-11 (b) image subtraction sections.** The frames show the difference images from the in-transit points where a diffraction spike is present. The candidate star coordinates are centred in the frames and each frame is 19.0" by 13.3".

Several other candidates were also eliminated on examination of the in-transit difference images. These candidates were found to have diffraction spikes in the in-transit points. These candidates are shown with lightcurves and in-transit difference images in Figures 5.10 to 5.14. Again the lightcurves are shown without error bars and with a finder image along side. Also shown are the in-transit difference images.

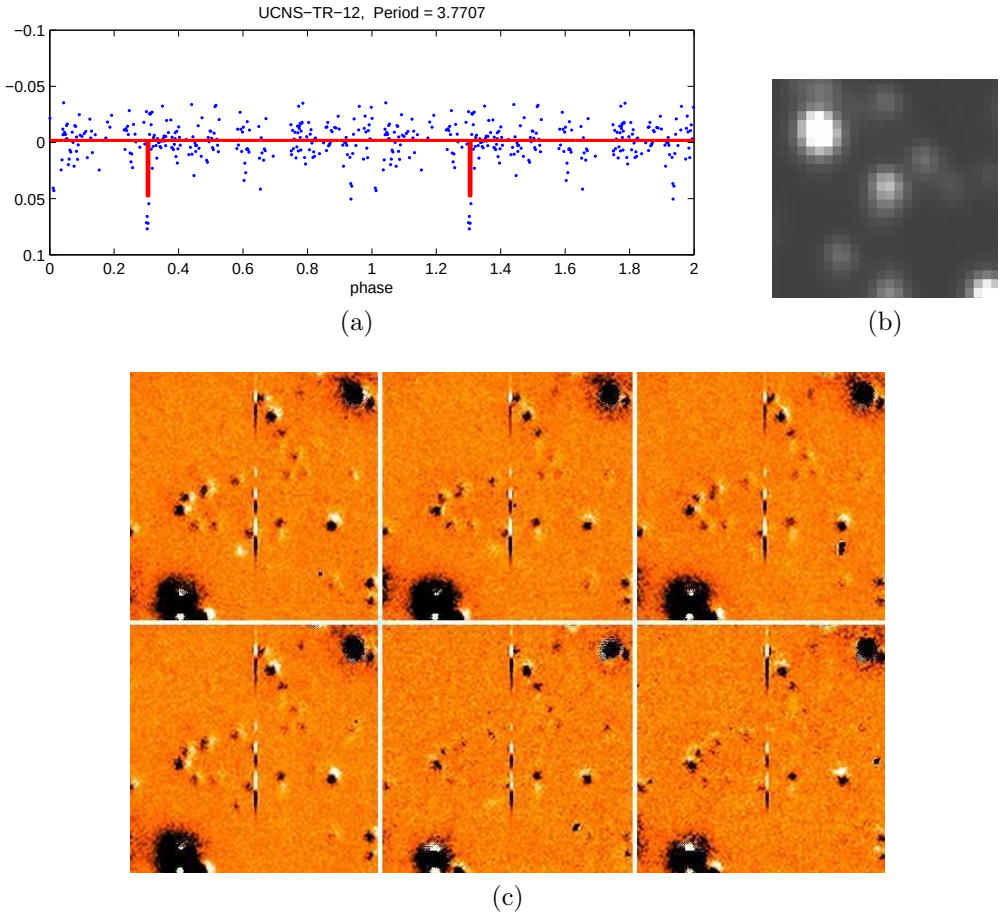


Figure 5.10: **Rejected candidate UCNS-TR-12**: transit candidate lightcurve rejected due to image. Figure (a) shows the initial de-trended reduction, figure (b) a finder image and figure (c) the in-transit difference images with diffraction spike (each of these frames is 26.2" by 26.2" with the candidate centred).

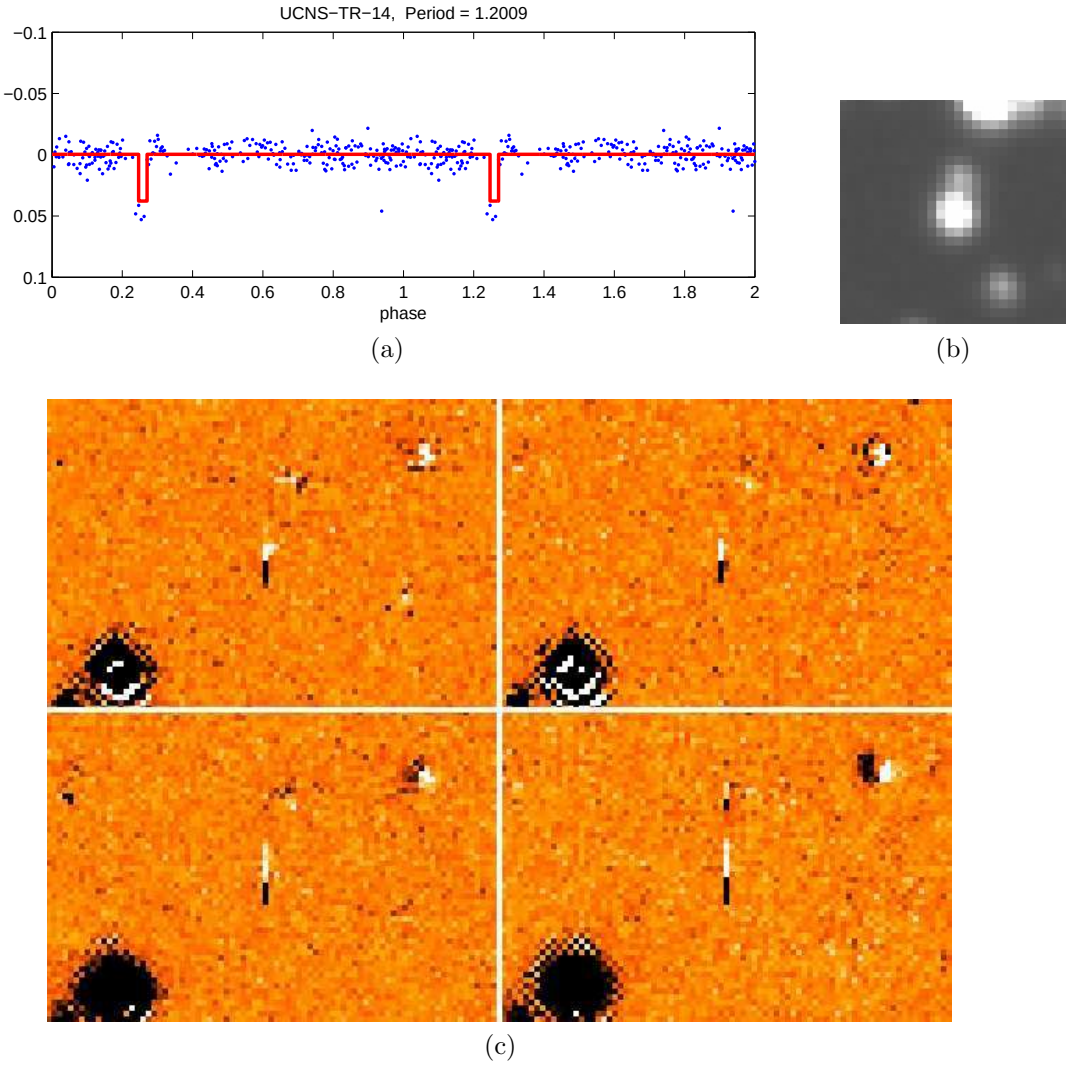
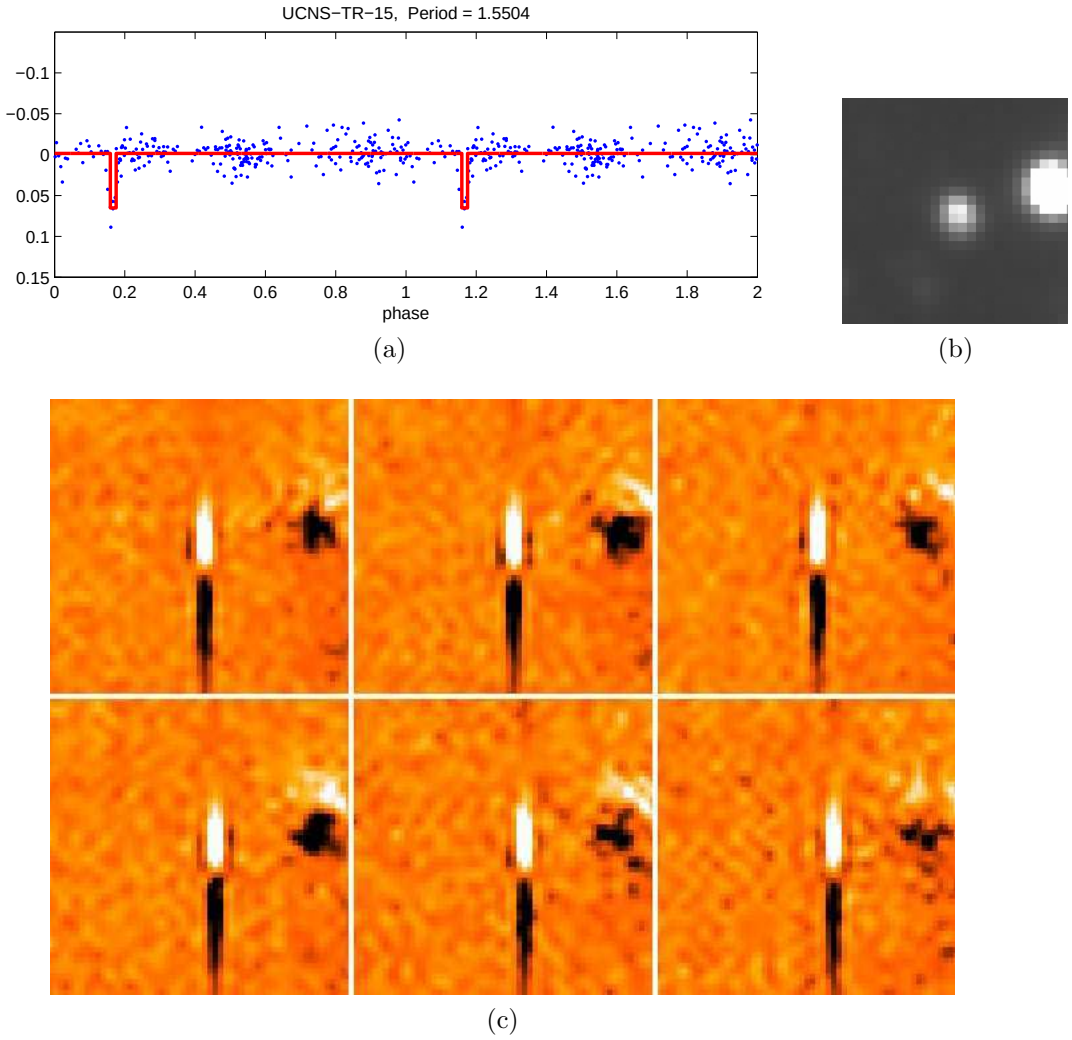


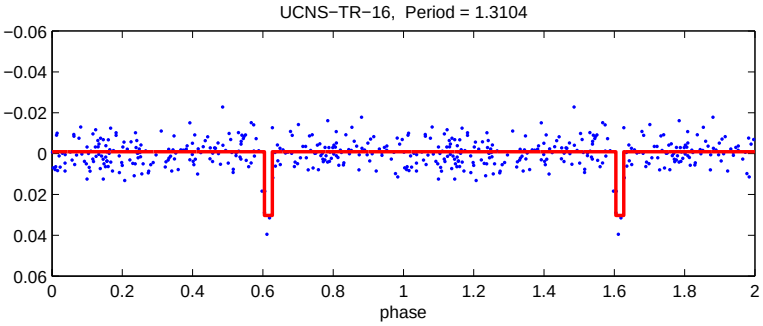
Figure 5.11: **Rejected candidate UCNS-TR-14**:transit candidate lightcurve rejected due to image. Figure (a) shows the initial de-trended reduction, figure (b) a finder image and figure (c) the in-transit difference images with diffraction spike (each of these frames is 18.9" by 12.9" with the candidate centred).



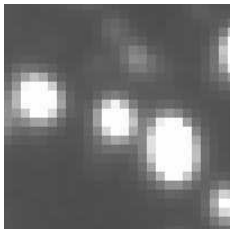
(a) (b)

(c)

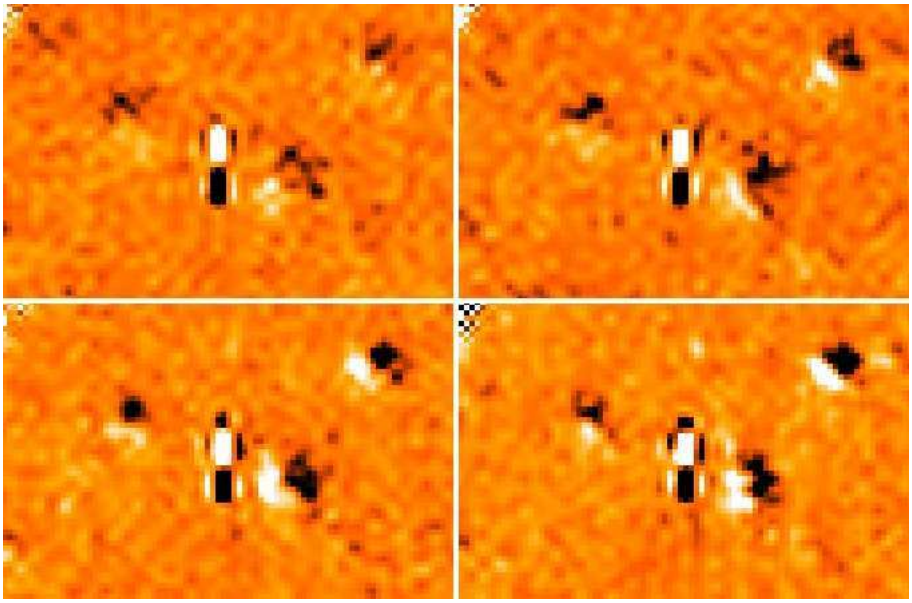
Figure 5.12: **Rejected candidate UCNS-TR-15**:transit candidate lightcurve rejected due to image. Figure (a) shows the initial de-trended reduction, figure (b) a finder image and figure (c) the in-transit difference images with diffraction spike (each of these frames is 6.7" by 6.7" with the candidate centred).



(a)



(b)



(c)

Figure 5.13: **Rejected candidate UCNS-TR-16**: transit candidate lightcurve rejected due to image. Figure (a) shows the initial de-trended reduction, figure (b) a finder image and figure (c) the in-transit difference images with diffraction spike (each of these frames is 9.9" by 13.3" with the candidate centred).

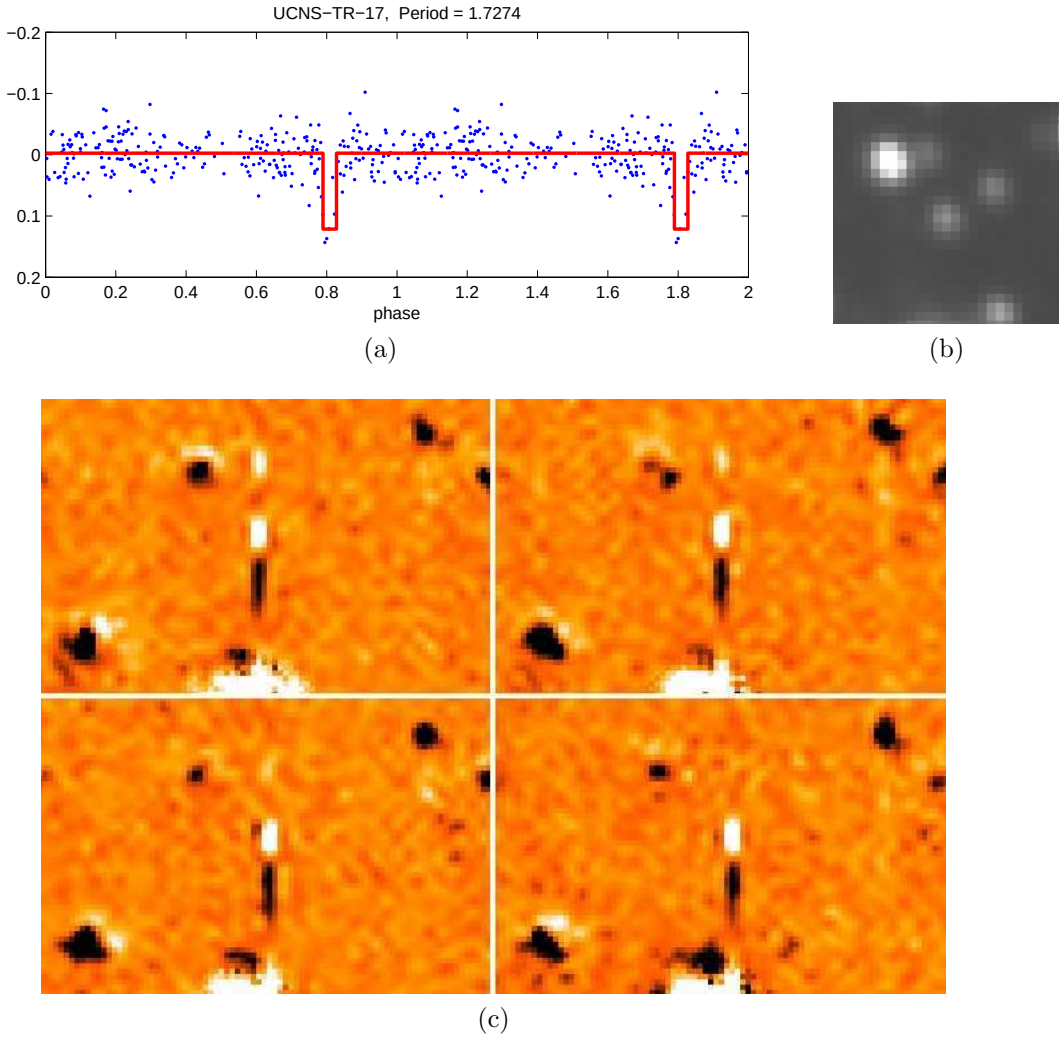
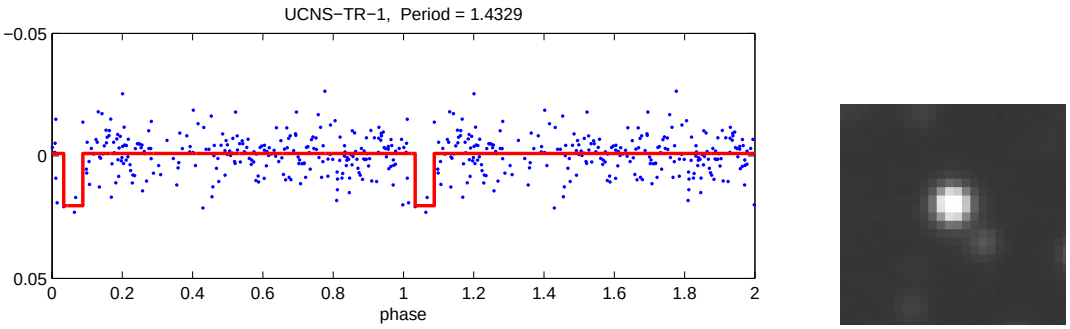


Figure 5.14: **Rejected candidate UCNS-TR-17**: transit candidate lightcurve rejected due to image. Figure (a) shows the initial de-trended reduction, figure (b) a finder image and figure (c) the in-transit difference images with diffraction spike (each of these frames is 10.0" by 6.7" with the candidate centred).

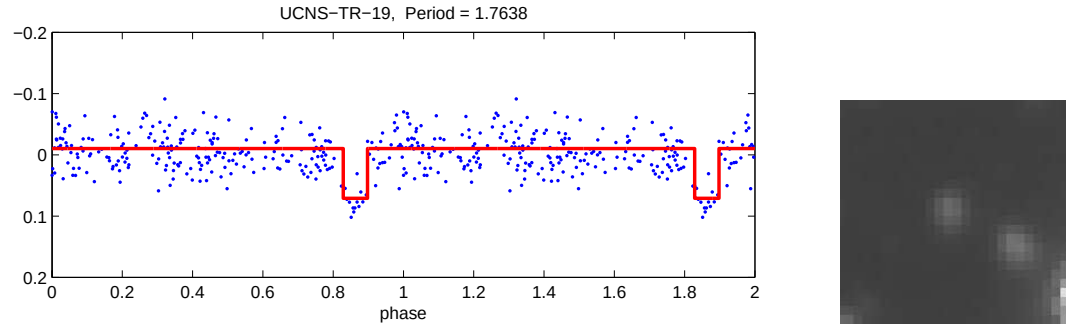
An eighth candidate UCNS-TR-1 was rejected due to extremely bad subtraction found in the region of the image where the candidate is located when examination of the in-transit points was performed. Examination of the in-transit points of candidate UCNS-TR-19 found that a number of other stars also had bright residuals in the in-transit difference images, consequently this candidate was rejected. Four other candidates have close neighbours or what may be a close neighbour; these candidates are not eliminated at this stage but will be closely examined after

re-reduction. The other of the candidates are clear of image effects.



(a) (b)

Figure 5.15: **Rejected candidate UCNS-TR-1**: transit candidate lightcurve rejected due to image.



(a) (b)

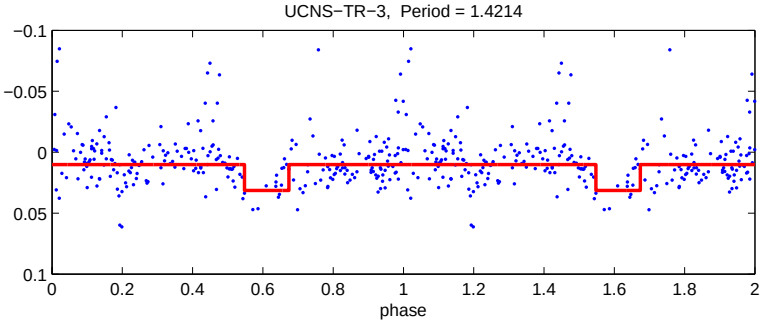
Figure 5.16: **Rejected candidate UCNS-TR-19**: transit candidate lightcurve rejected due to image.

While it was disappointing to have to reject these it gives good confidence that the search algorithm is detecting transit-like lightcurves.

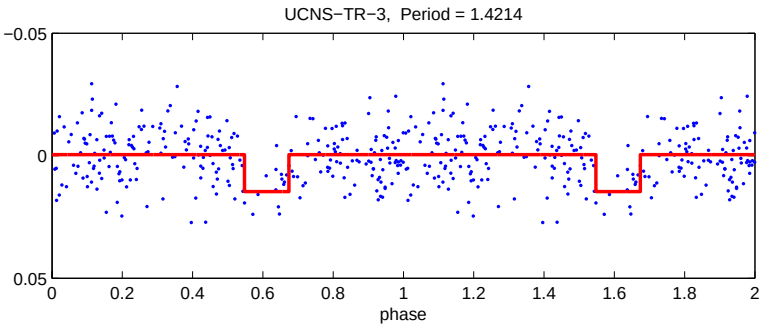
5.7 Re-reduction of candidates

After examination of the images is completed the list of candidates passing this test were re-reduced to determine if the signal is caused by defects in the photometry. The reduction is performed by PySIS3 (see §3.1.5) in the form it used for microlensing events. This reduces the time involved in the reduction. The subtraction was performed on small sections of the images surrounding the candidate coordinates. The photometry was performed on the candidate star and the resulting lightcurve examined. If the transit signal is apparent in the new lightcurve the star is considered further as a planetary transit candidate. Light curves of the candidates which no longer show the transit signal are rejected during this process. The filtered, de-trended and re-reduced lightcurves are shown phased by detected period in Figures 5.17 to 5.38, and again the lightcurves are plotted here without the error bars for ease of viewing. The solid line shows the detected transit signal where the depth is calculated from the mean of out-of-transit points - mean of in-transit points. This calculation is performed separately for each of the three lightcurves. In a number of cases the signal detected in the de-trended lightcurves can also be seen in the un-de-trended lightcurves. Also shown are the un-phased original lightcurves, in separate figures, with the determined transit signal as a solid line.

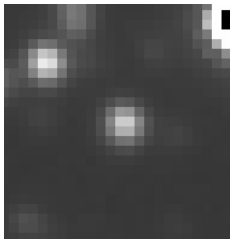
Candidate UCNS-TR-3 did not have a strong transit signal in the initial reduction, Figure 5.17(b). There is no sign of a transit-like shape in the lightcurve from the re-reduction, Figure 5.17(d). Examination of the un-phased lightcurve, Figure 5.18 shows no real signal.



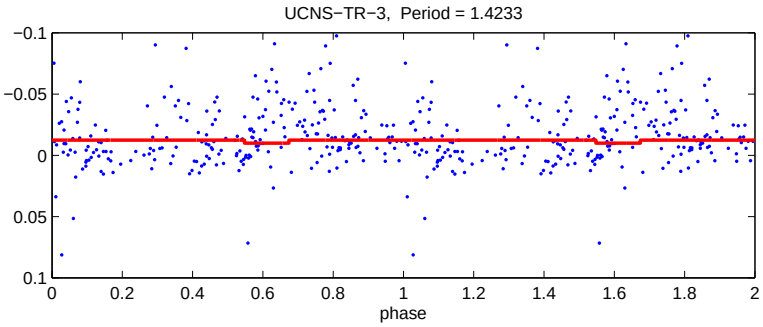
(a)



(b)



(c)



(d)

Figure 5.17: **Rejected candidate UCNS-TR-3**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

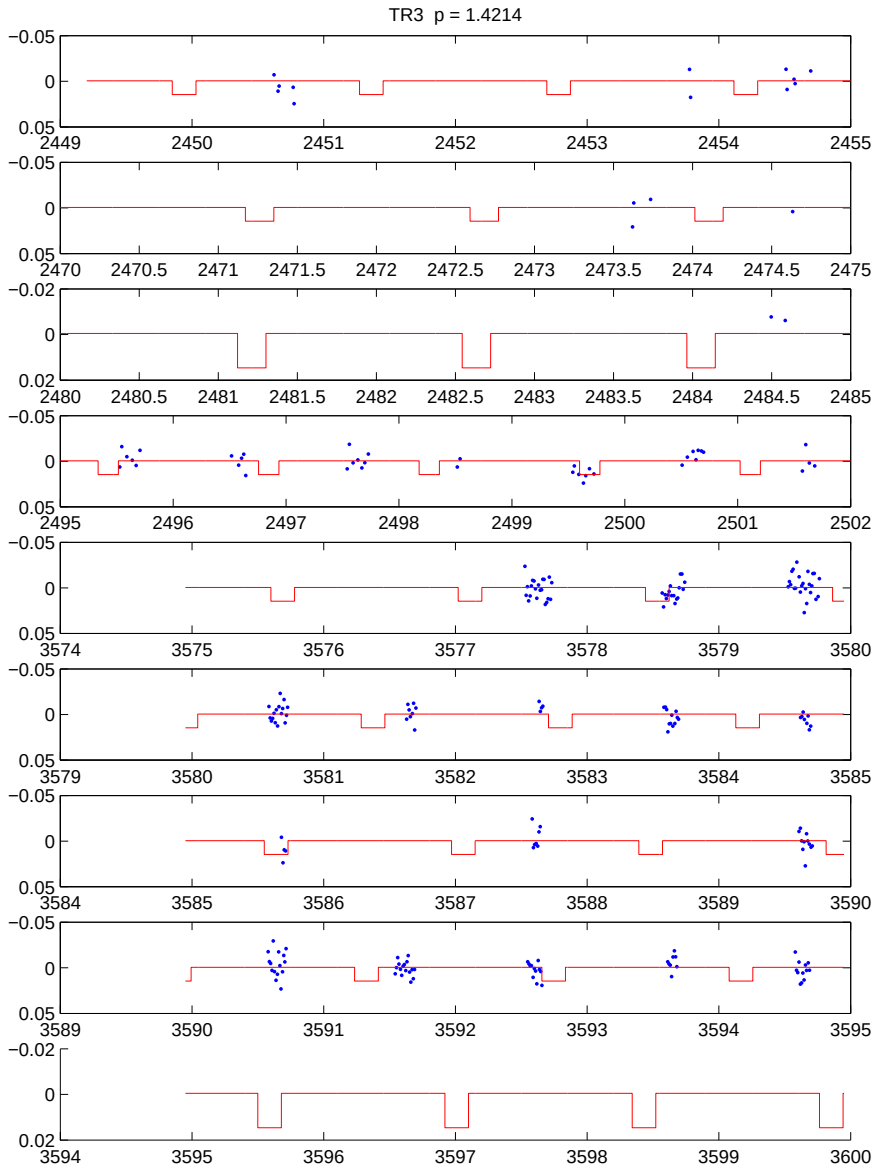


Figure 5.18: **Un-phased candidate UCNS-TR-3** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

Candidate UCNS-TR-5 had a strong narrow signal in the original reduction, Figure 5.19(b). The lightcurve resulting from the re-reduction has no signal of decreased magnitude but rather a shallow increase in magnitude at the same period, Figure 5.19(d). The in-transit points in the un-phased lightcurve, Figure 5.20, are not significantly different from those surrounding.

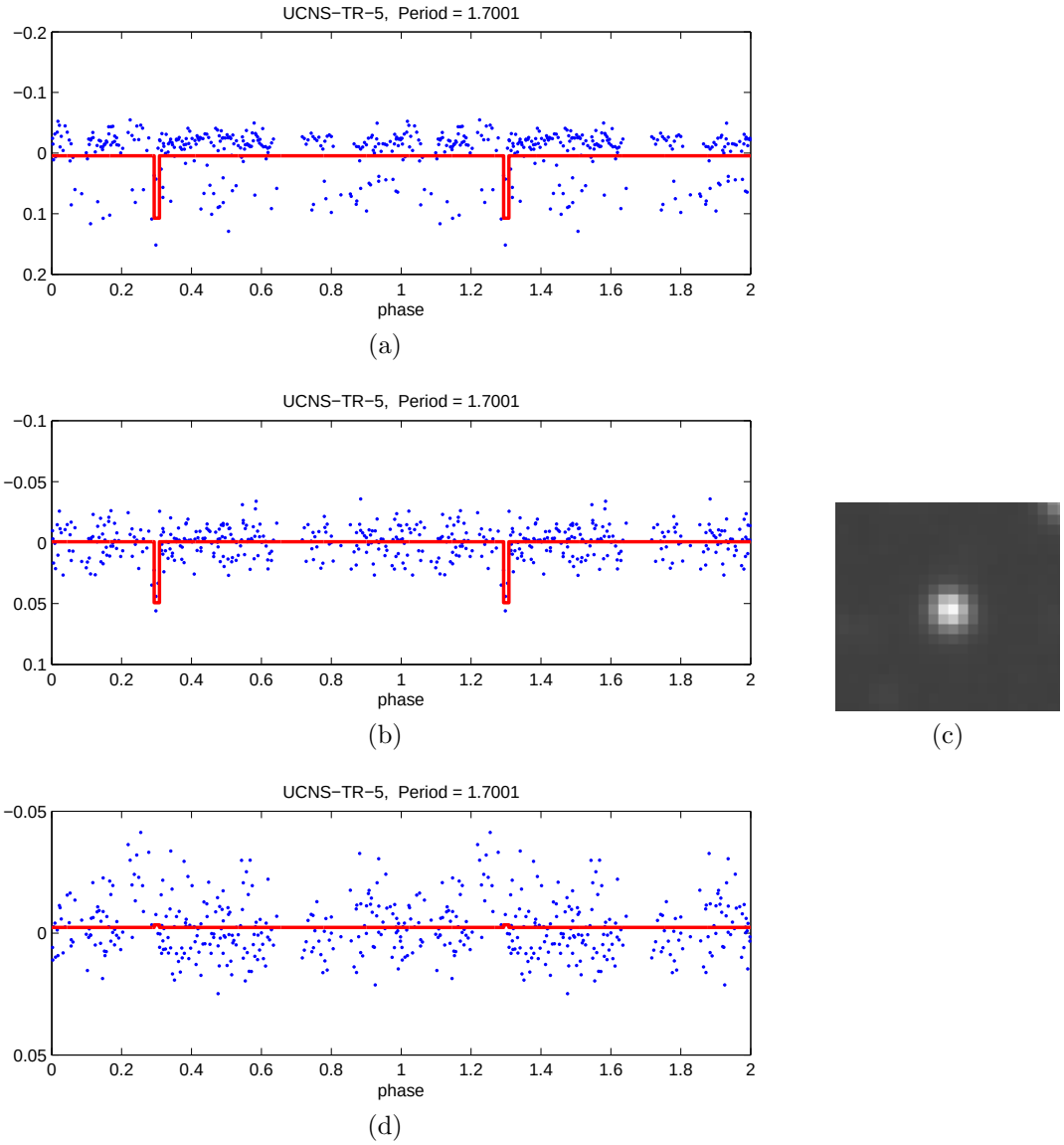


Figure 5.19: **Rejected candidate UCNS-TR-5**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

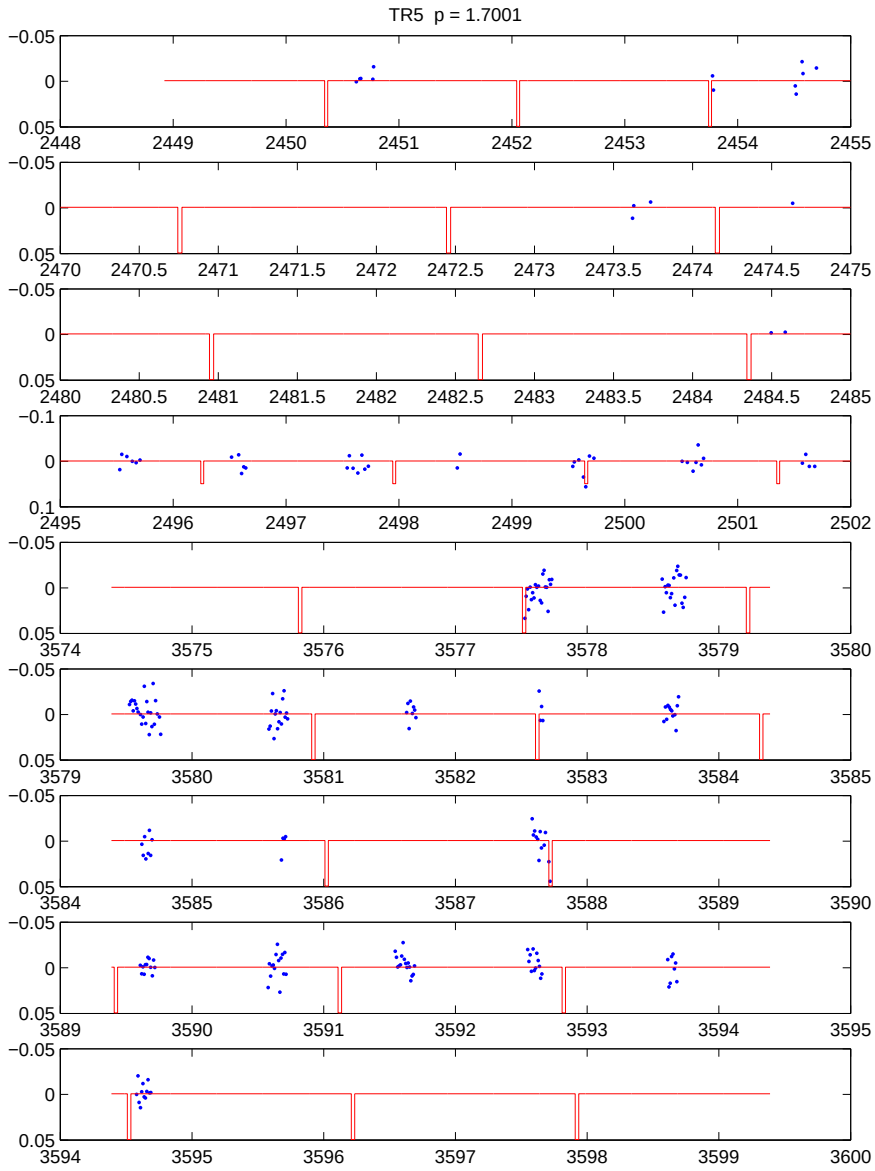


Figure 5.20: **Un-phased candidate UCNS-TR-5** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

Candidate UCNS-TR-6 had a weak signal (due to noise) in the original reduction, Figure 5.21(b). Examination of the lightcurve from the re-reduction shows an increased amount of noise making a transit signal difficult to identify, Figure 5.21(d). From the un-phased lightcurve, Figure 5.22, it appears that the transit signal is caused by a single cluster of points not significantly different from any others.

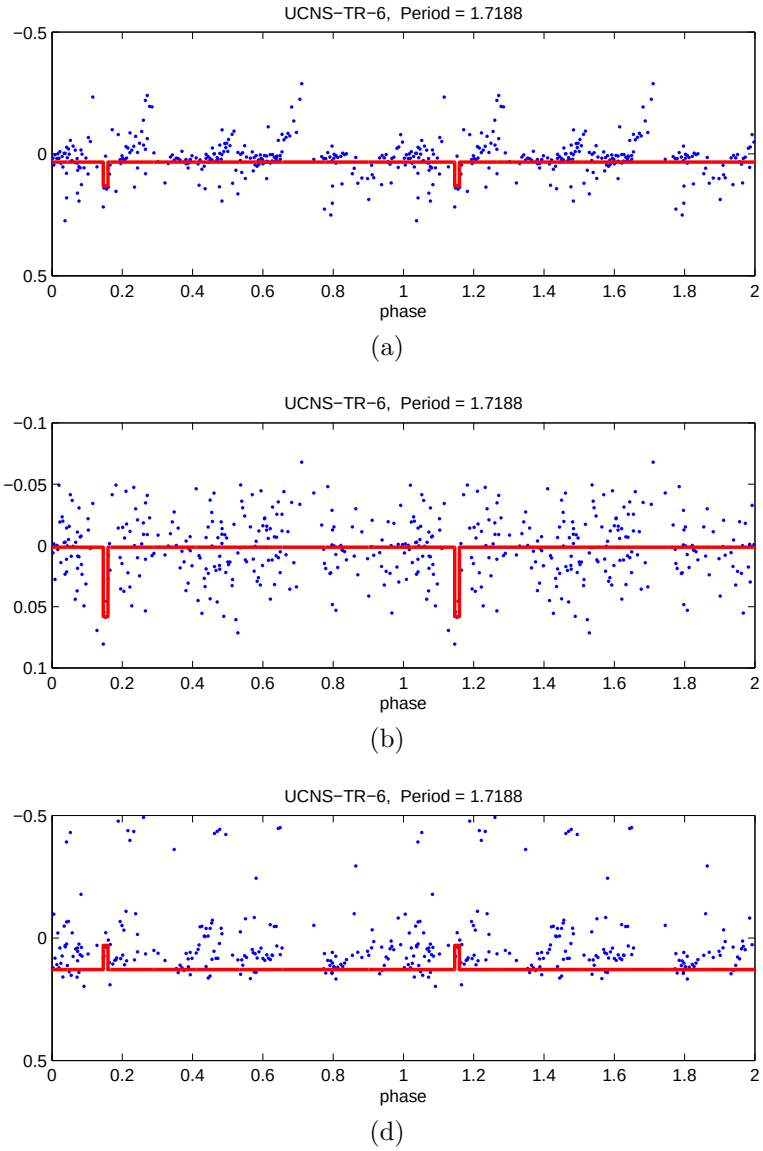


Figure 5.21: **Rejected candidate UCNS-TR-6**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

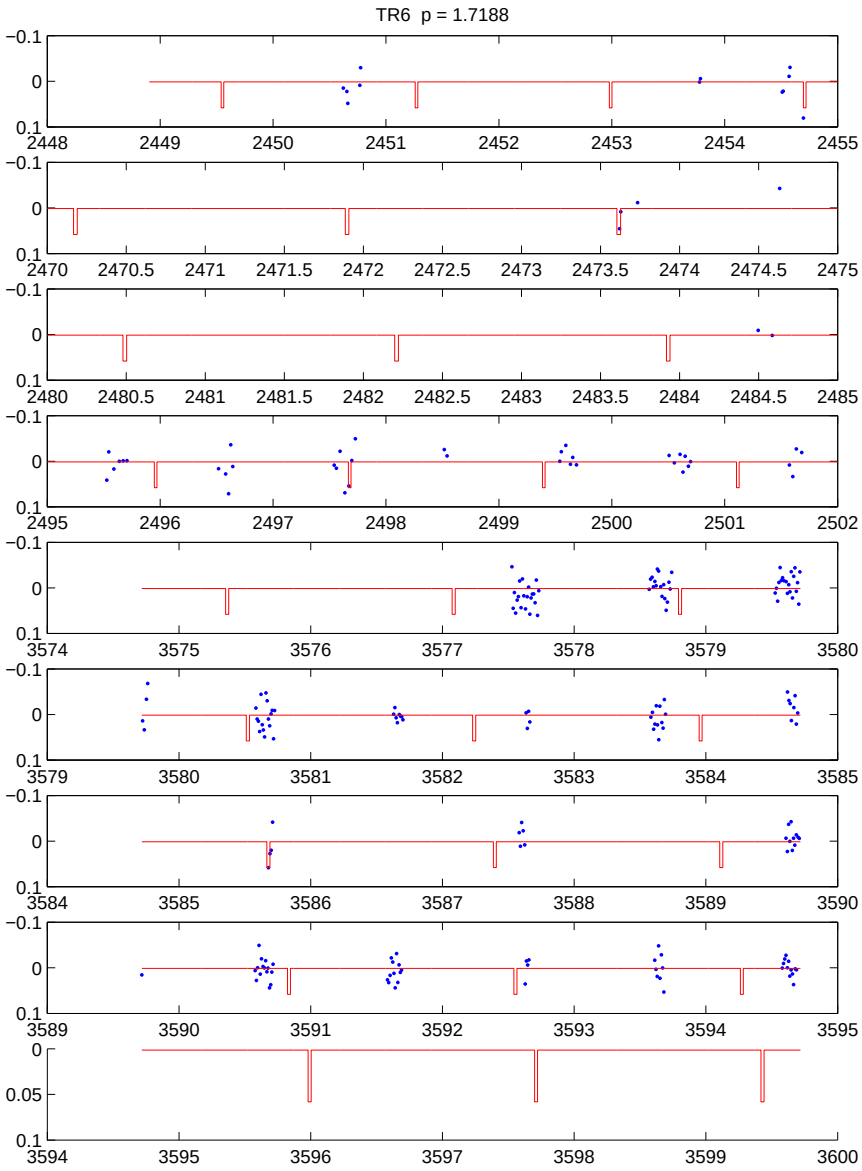


Figure 5.22: **Un-phased candidate UCNS-TR-6** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

Candidate UCNS-TR-7 was not a good candidate even in the original reduction with a narrow signal in a noisy light curve, Figure 5.23(b). None of this signal has persisted into the re-reduction, Figure 5.23(d). The un-phased lightcurve, Figure 5.24, shows several cycles with points following the transit well. Despite the lack of signal in the re-reduced (un-de-trended) data this should be further considered.

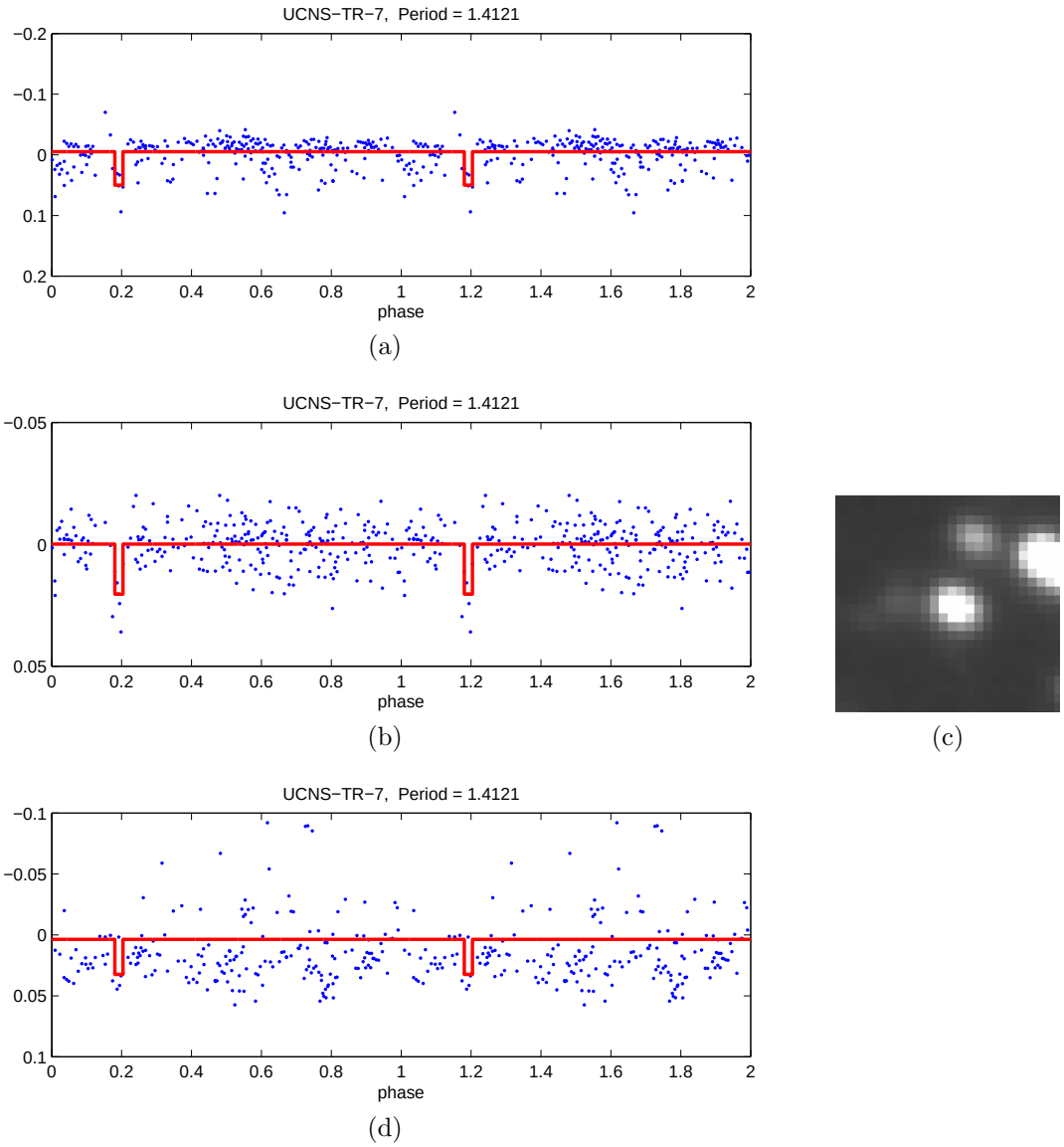


Figure 5.23: **Rejected candidate UCNS-TR-7**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

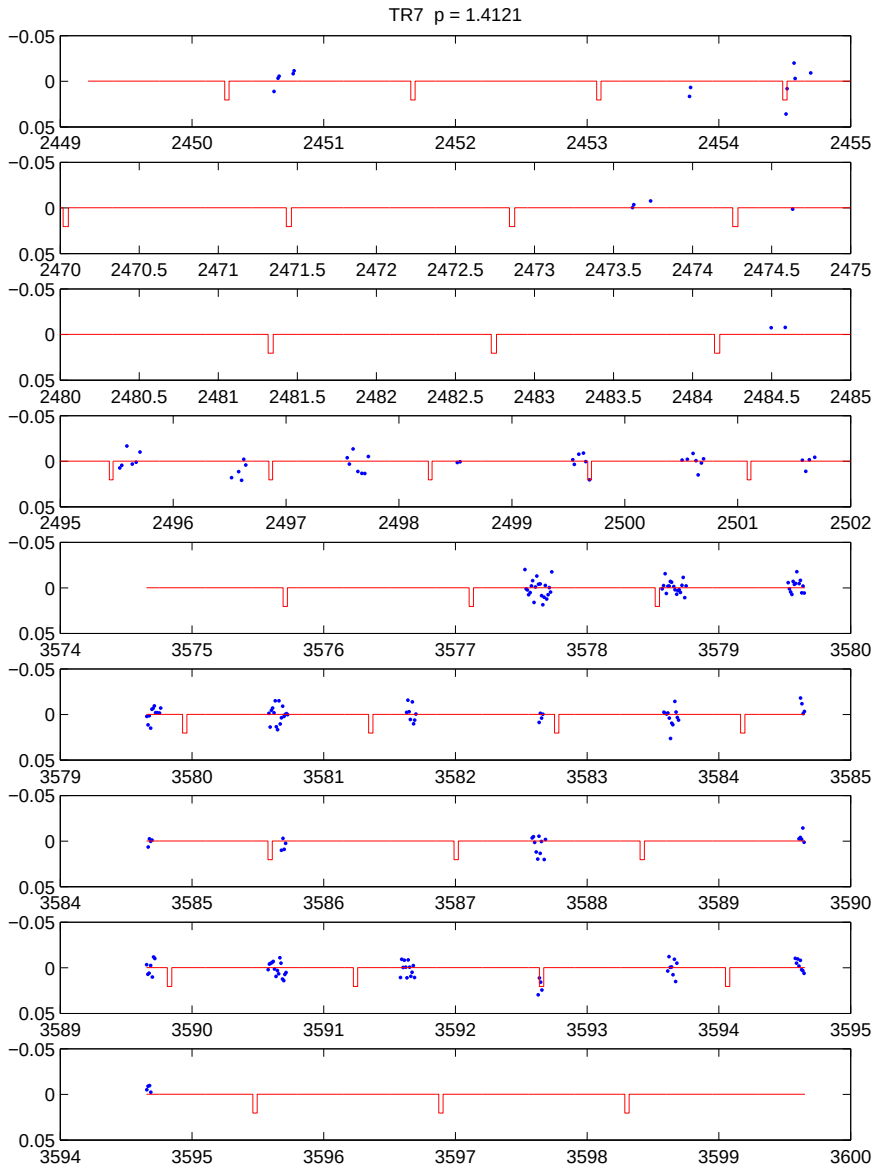
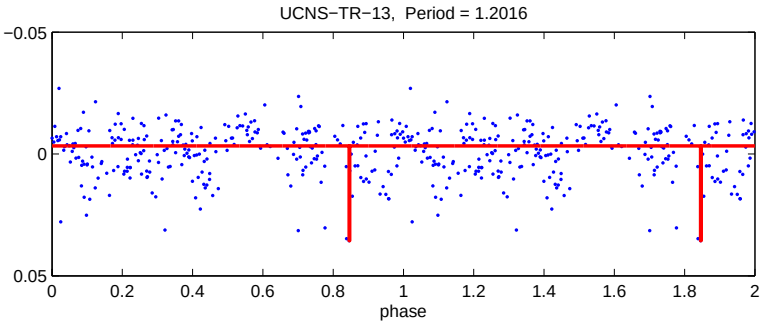
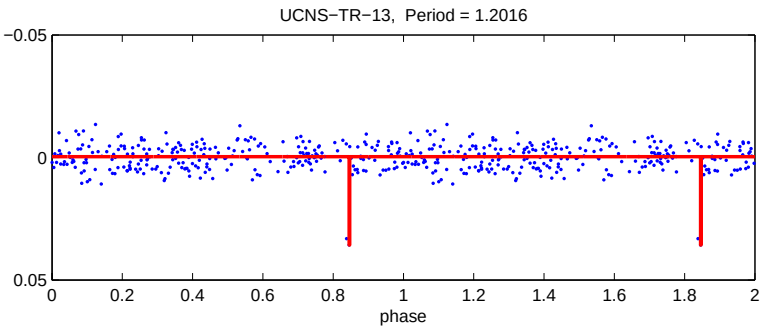


Figure 5.24: **Un-phased candidate UCNS-TR-7** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

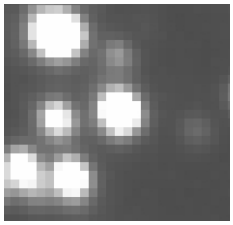
Candidate UCNS-TR-13 had only 2 points indicating a possible transit in the initial reduction, Figure 5.25(b). In the re-reduction a large amount of scatter and noise was found in the lightcurve, Figure 5.25(d). The un-phased lightcurve only shows a single cycle with in-transit points, Figure 5.26, however this is the only lightcurve found with continuous points through a transit well in the un-phased lightcurve. This candidate should be further considered.



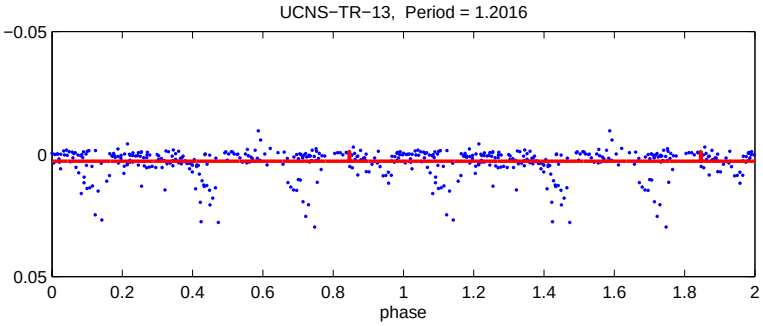
(a)



(b)



(c)



(d)

Figure 5.25: **Rejected candidate UCNS-TR-13**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

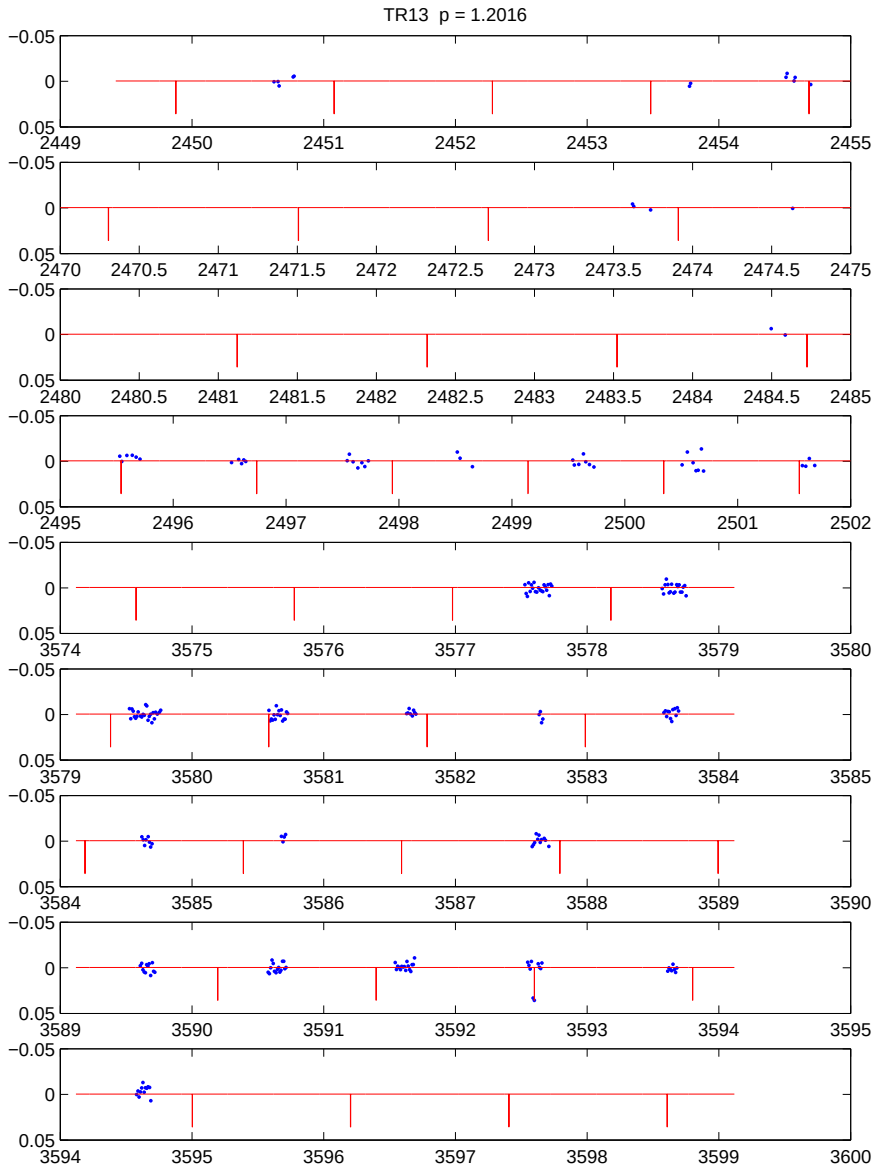
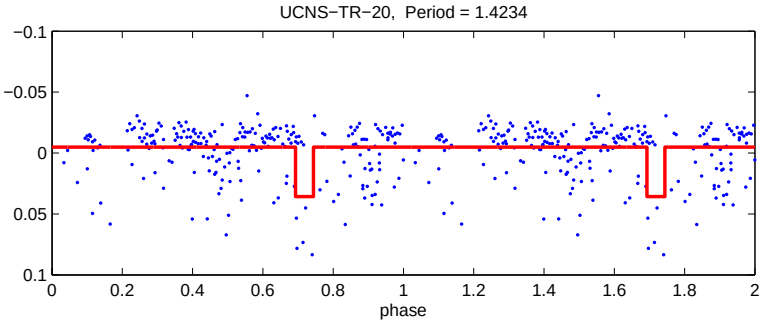
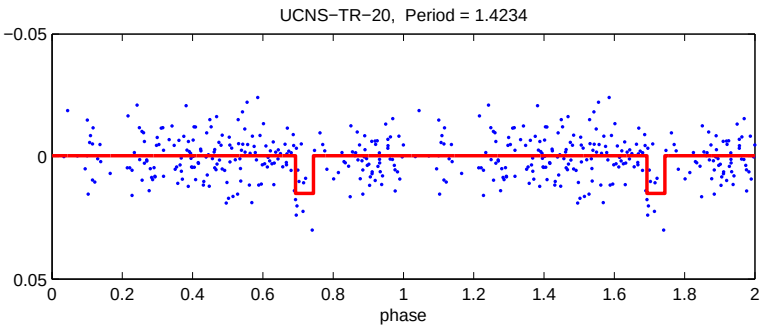


Figure 5.26: **Un-phased candidate UCNS-TR-13** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

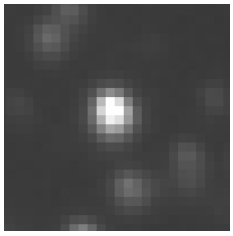
The minimal transit signal of candidate UCNS-TR-20 seen in the original lightcurve, Figure 5.27(b) is not visible in the re-reduction, Figure 5.27(d). The un-phased lightcurve, Figure 5.28, shows large scatter and insignificant deviations in the transit wells.



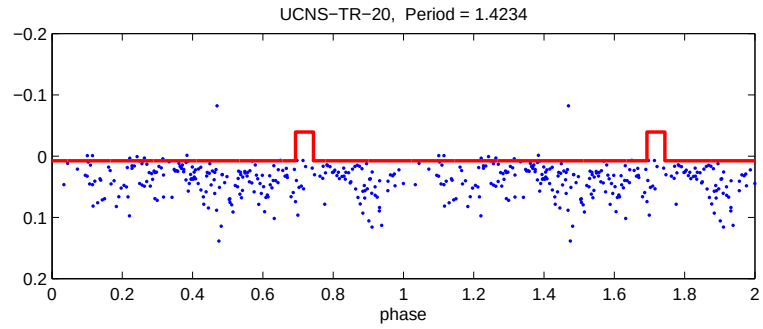
(a)



(b)



(c)



(d)

Figure 5.27: **Rejected candidate UCNS-TR-20**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

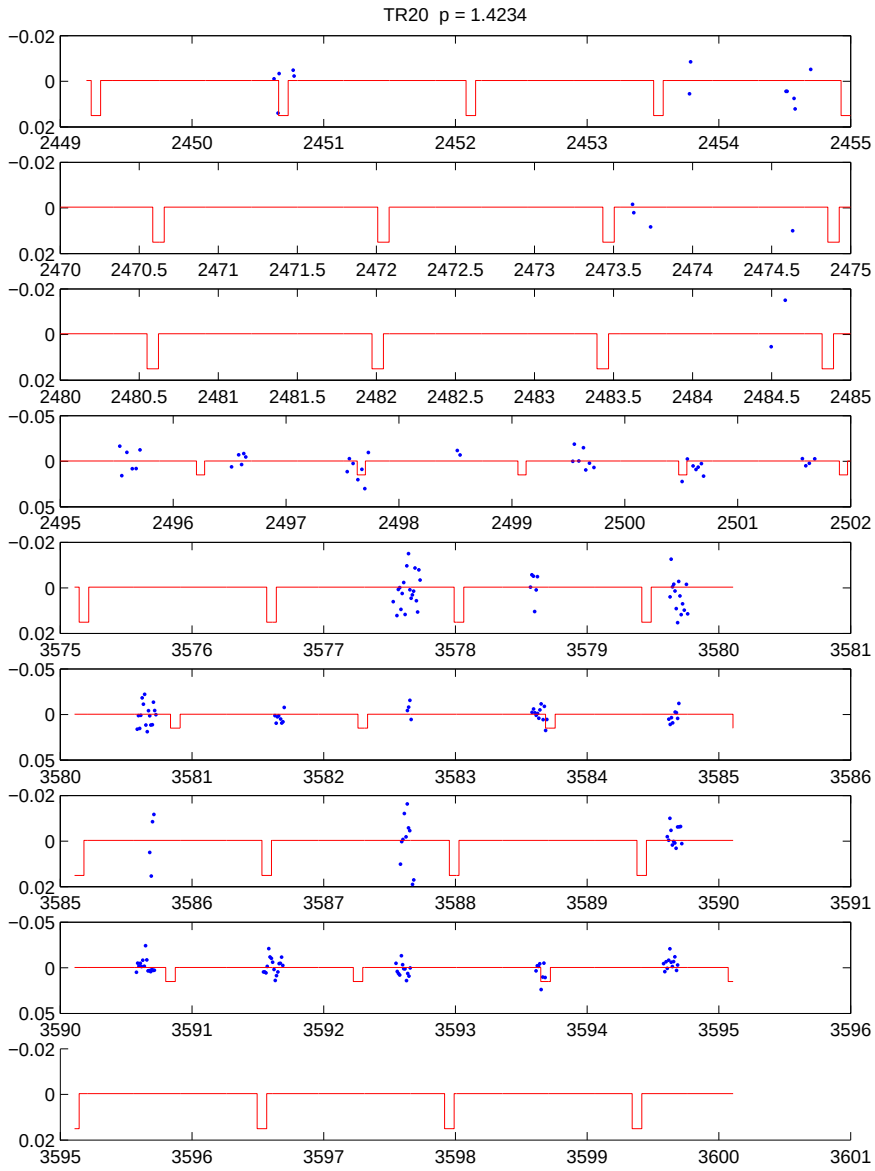


Figure 5.28: **Un-phased candidate UCNS-TR-20** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

The original un-de-trended lightcurve of UCNS-TR-23, Figure 5.29(a), appears to be varying with a very low amplitude, but this was not detected in the variable search. The transit signal of candidate UCNS-TR-23 is very weak in the original lightcurve, Figure 5.29(b). The re-reduction appears much noisier and the transit signal is not visible, Figure 5.29(d). Despite the length of this signal, no points fall across the bottom of the well as seen in the un-phased lightcurve, Figure 5.30.

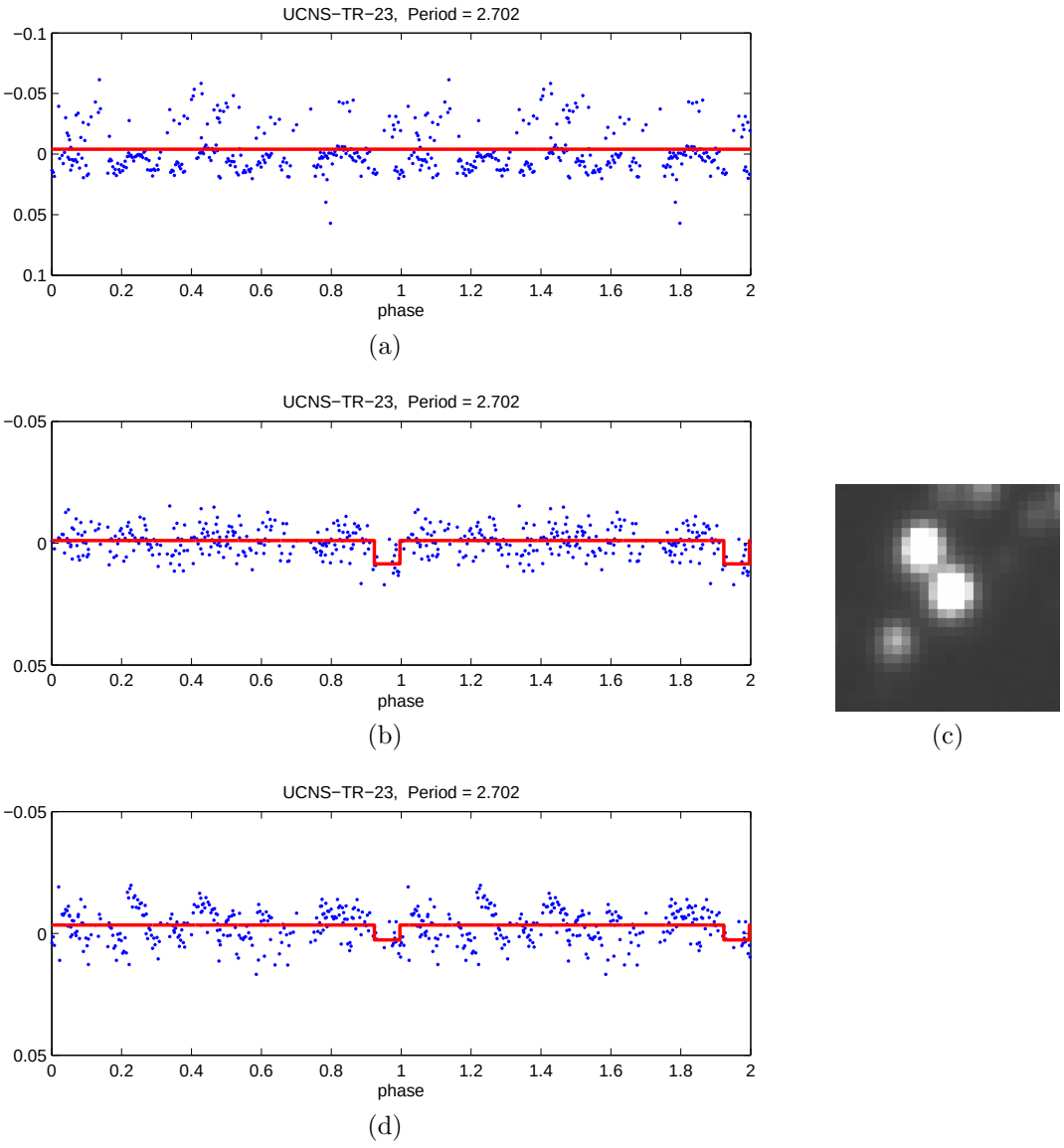


Figure 5.29: **Rejected candidate UCNS-TR-23**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

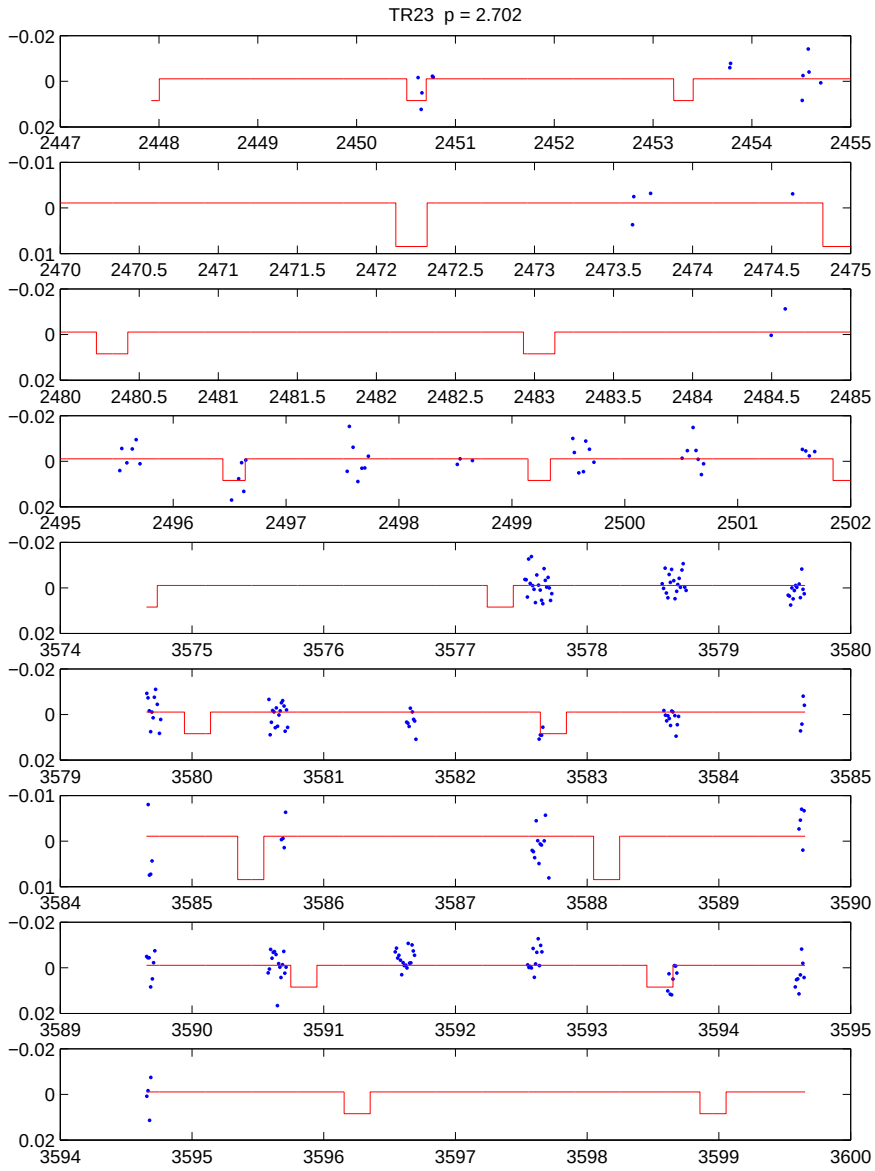
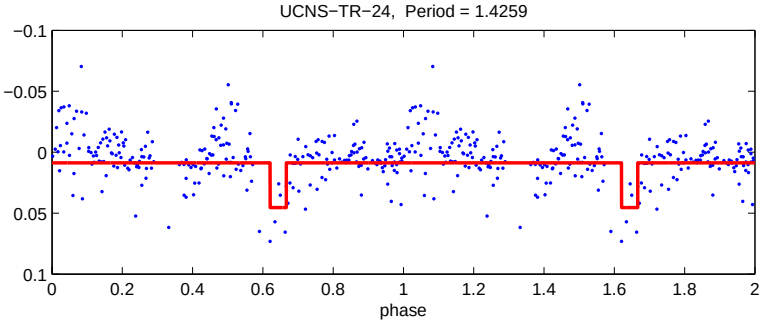
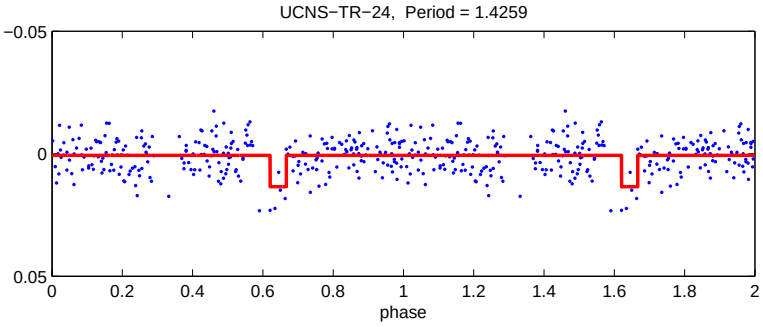


Figure 5.30: **Un-phased candidate UCNS-TR-23** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

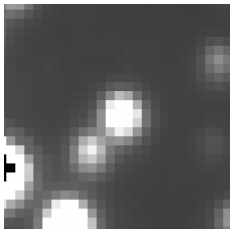
Candidate UCNS-TR-24 contained only a few points in-transit in the original lightcurve, Figure 5.31(b). The lightcurve produced by the re-reduction has no similar shape, Figure 5.31(d). The level of nightly scatter visible in the un-phased lightcurve, Figure 5.32, makes this detection unconvincing.



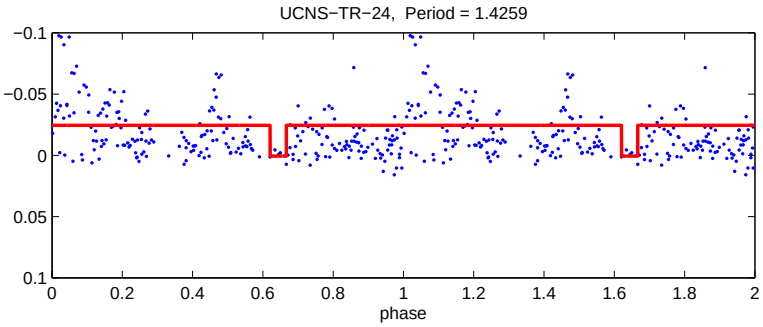
(a)



(b)



(c)



(d)

Figure 5.31: **Rejected candidate UCNS-TR-24**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

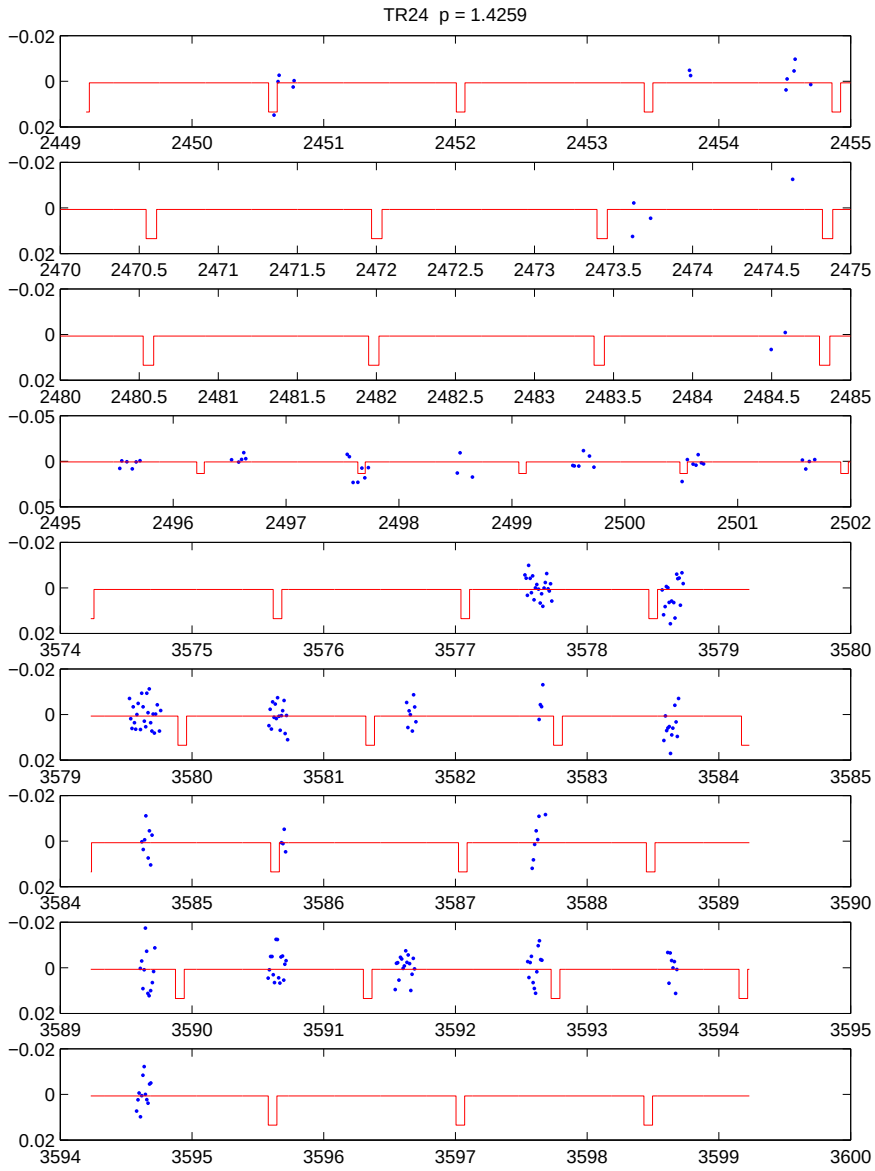
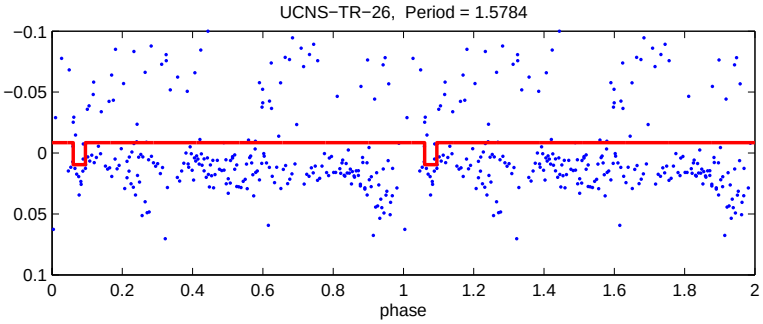
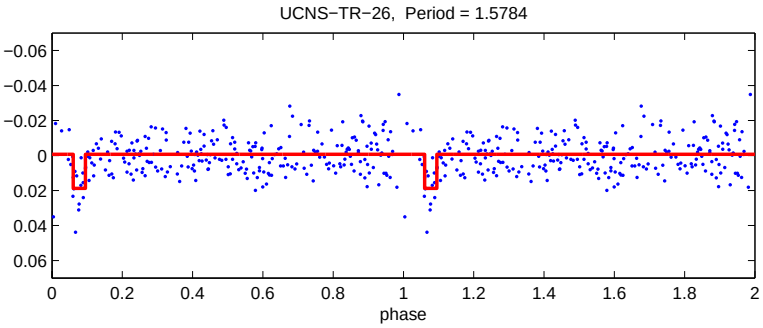


Figure 5.32: **Un-phased candidate UCNS-TR-24** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

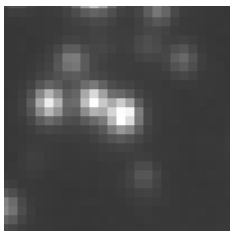
The original lightcurve, Figure 5.33(b) of UCNS-TR-26 showed a reasonable candidate. The re-reduction shows two stronger dips in the lightcurve - though re-computing the period made no improvement - and the lightcurve as a whole appears an unlikely candidate, Figure 5.33(d). The un-phased lightcurve, Figure 5.34, shows three cycles with in-transit points however the scatter amongst the other points indicates this is most likely coincidental.



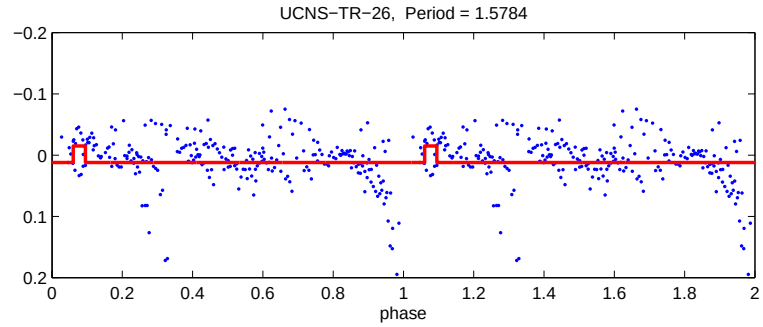
(a)



(b)



(c)



(d)

Figure 5.33: **Rejected candidate UCNS-TR-26**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

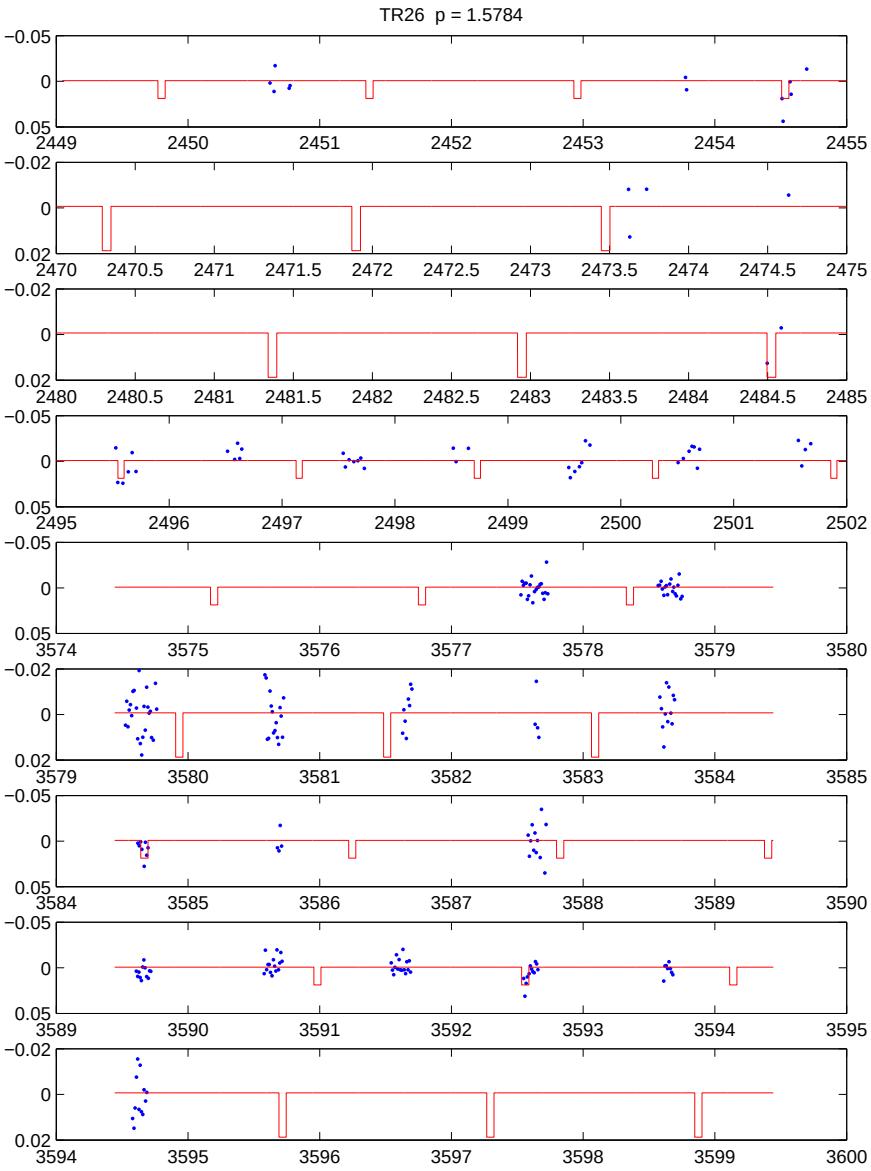
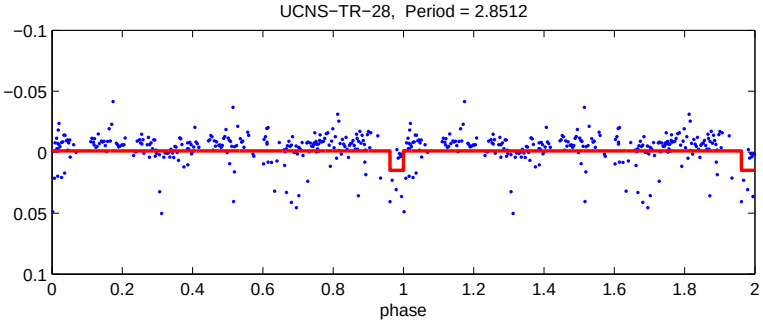
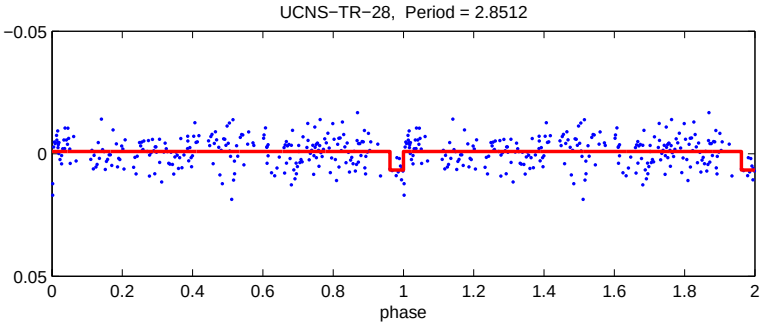


Figure 5.34: **Un-phased candidate UCNS-TR-26** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

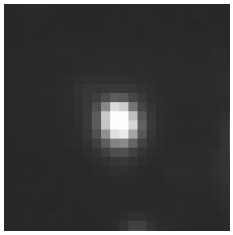
Candidate UCNS-TR-28 has a scattered, transit signal in a noisy original lightcurve, Figure 5.35(b). This signal does not carry through in the re-reduced lightcurve, Figure 5.35(d). The scatter in the un-phased lightcurve, Figure 5.36, completely overwhelms the single cycle with in-transit points.



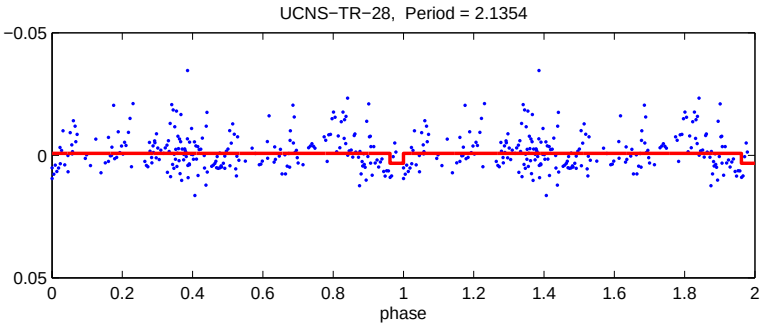
(a)



(b)



(c)



(d)

Figure 5.35: **Rejected candidate UCNS-TR-28**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

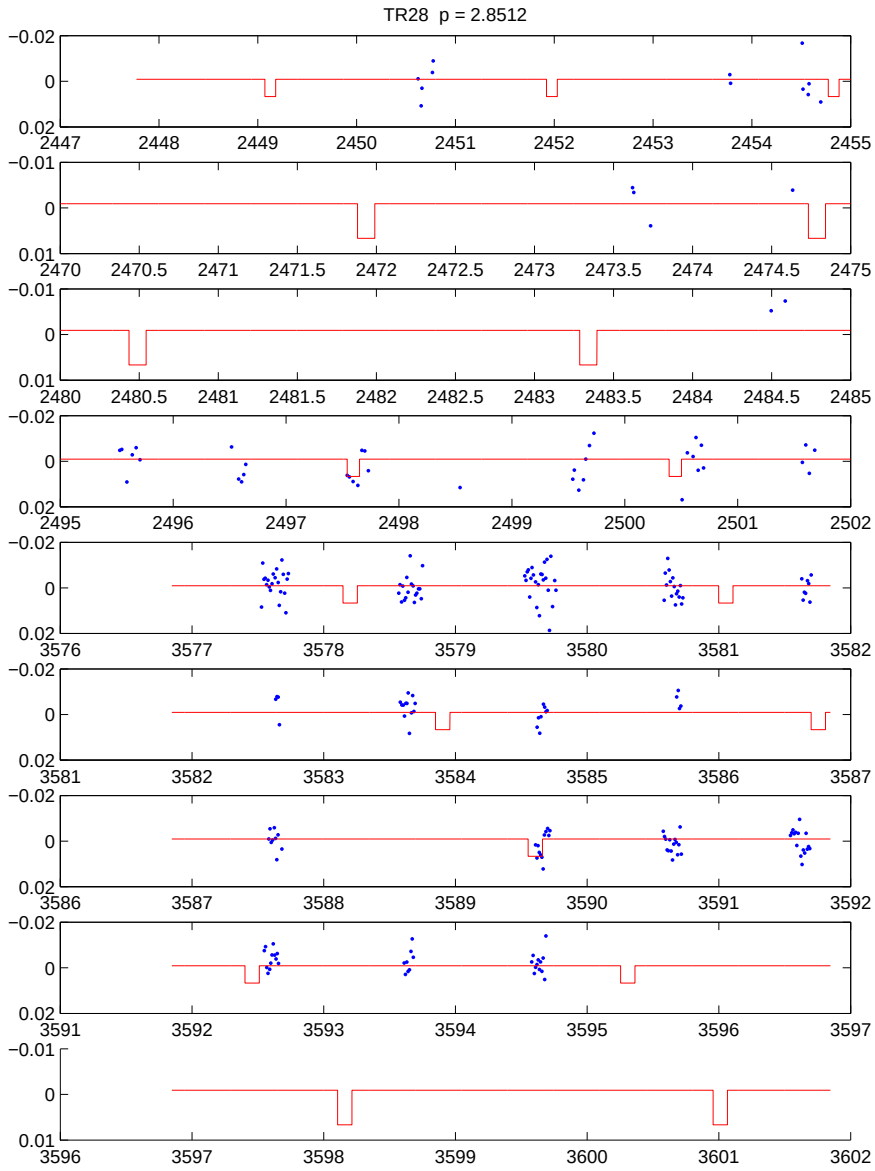


Figure 5.36: **Un-phased candidate UCNS-TR-28** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

The original lightcurve of candidate UCNS-TR-30 contains a very narrow dip, Figure 5.37(b). This is not reproduced in the re-reduced lightcurve, Figure 5.37(d). It seems likely that the original dip is caused either by the original photometry, the de-trending or indeed one of the close neighbours, Figure 5.37(b). The narrow signal detected in this lightcurve appears to have picked up only a few points, the un-phased lightcurve, Figure 5.38, shows a number of other points of similar depth not contained in the transit well.

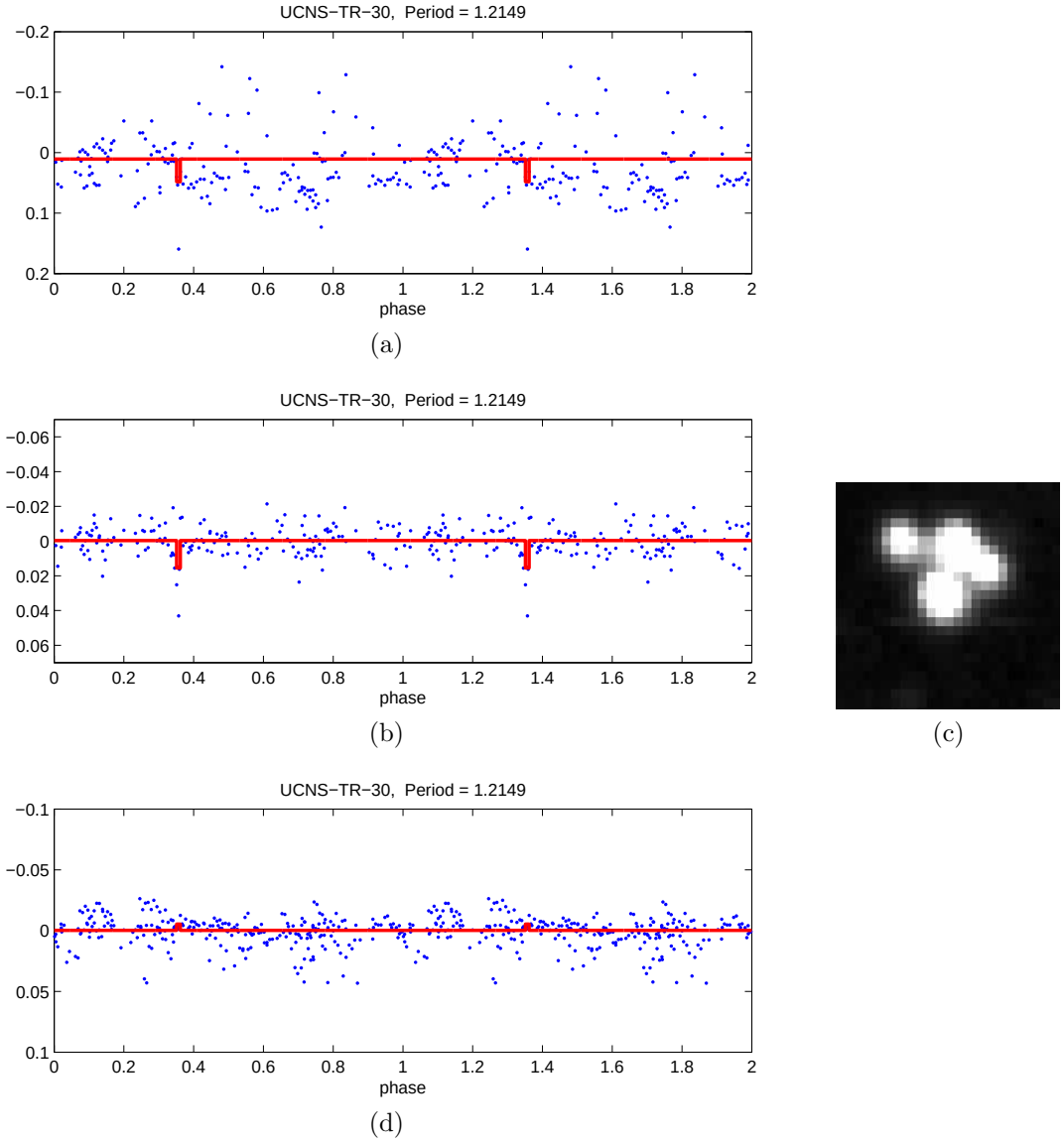


Figure 5.37: **Rejected candidate UCNS-TR-30**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

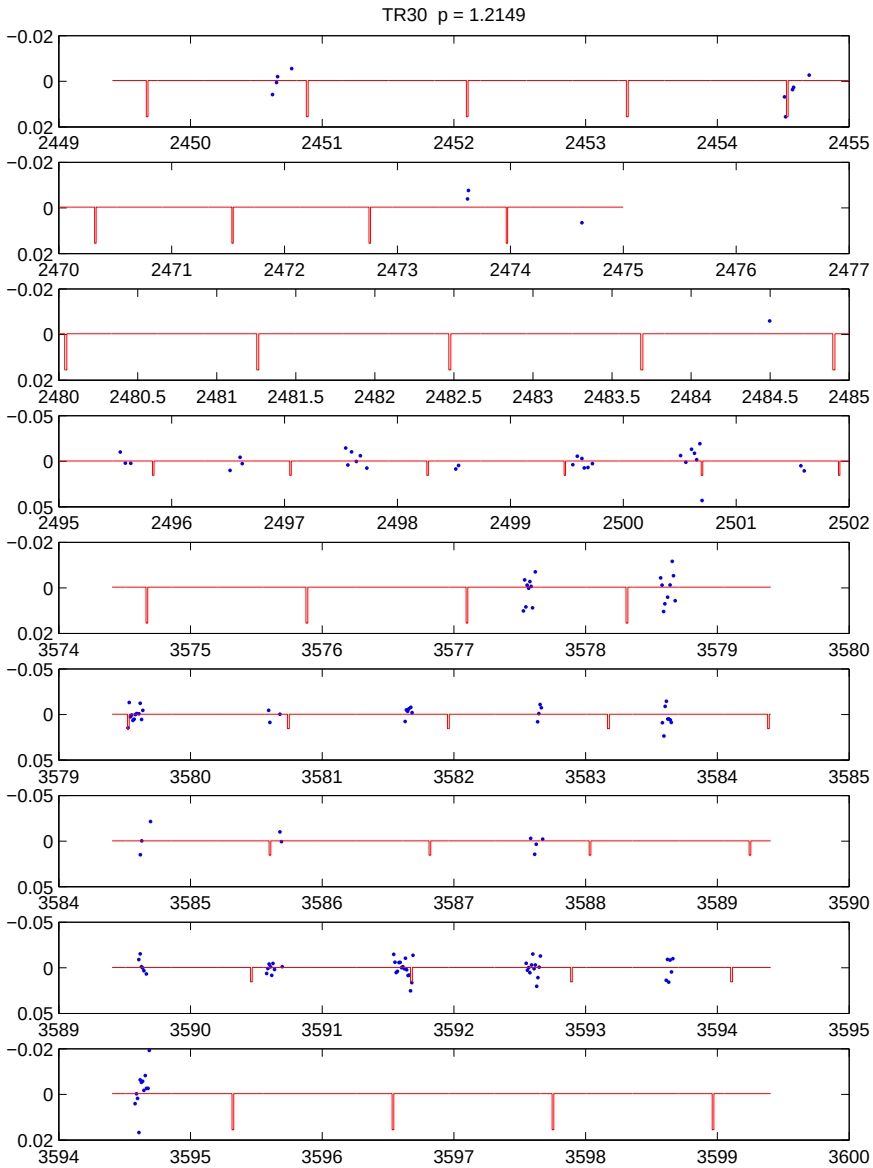


Figure 5.38: **Un-phased candidate UCNS-TR-30** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

Following the re-reduction of the transit candidates we are left with only three candidates which still show a transit like signature in their lightcurves. These candidates are discussed in the following section.

5.8 Remaining candidates

The three candidates remaining are recorded with parameters in Table 5.3. We summarize these candidates in need of further examination following. In all the following figures of candidates the lightcurves are plotted without error bars for ease of viewing and with a solid line showing the transit where the depth is calculated for each lightcurve.

ID	Period (days)	Phase	Epoch HJD - 2450000.0	Transit Length (hours)	Depth (mmag)	Calibrated magnitude (R)	(B-R)	(R-I)
UCNS-TR-2	1.2840	0.2210	4494.2838	0.72	77 ± 11	19.43	2.34	0.80
UCNS-TR-4	1.9224	0.6626	4615.0338	2.17	82 ± 9	19.18	2.59	0.56
UCNS-TR-31	2.6269	0.4025	4729.4772	4.28	71 ± 8	19.85	2.25	0.88

Table 5.3: **Parameters of good transit candidates:** tabulated are id, period, phase, transit length and depth. The period, phase and transit length result from the search algorithm, and phase indicates the start of the transit. The epoch shown is the heliocentric Julian date of the start of a specific transit and is defined by: 2450000.0 + (n + phase)*period where n is an integer. The depth of transit is calculated from the difference in weighted mean in and out of transit and here is shown for the original de-trended lightcurve.

Each good transit candidate has also been plotted on the appropriate colour magnitude diagram. All three candidates fall on the main sequence.

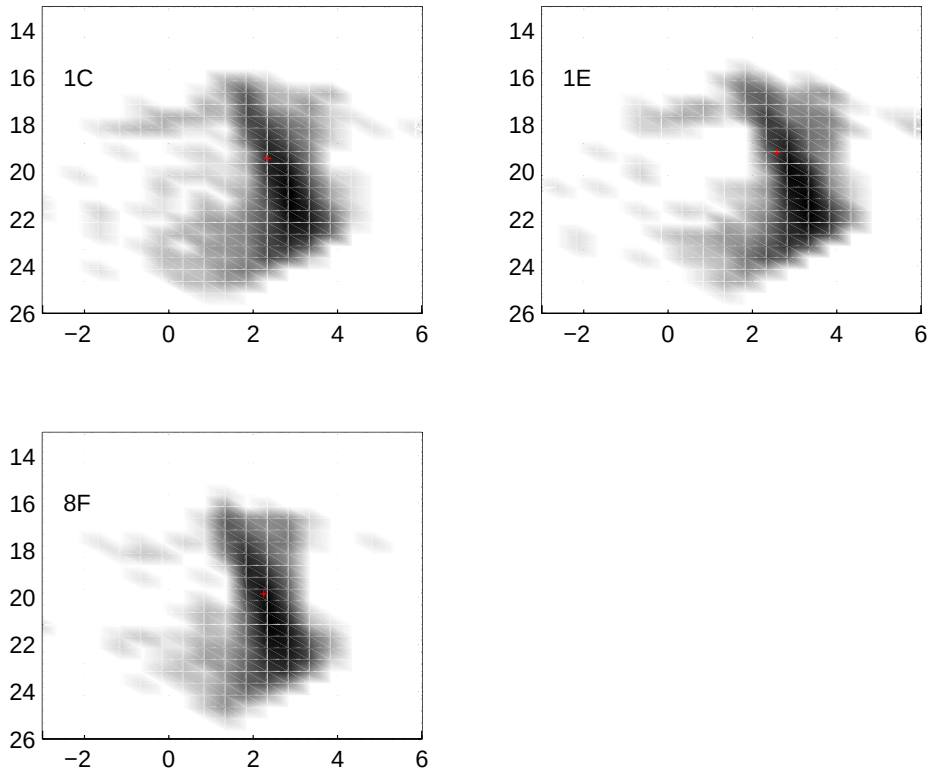


Figure 5.39: **Transit candidates on the colour-magnitude diagram** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the transit candidate from that sub-image over-plotted with +.

5.8.1 UCNS-TR-2

Candidate UCNS-TR-2 has a relatively deep transit signature (77 ± 11 mmag), Figure 5.41(b), which persists in the re-reduction lightcurve, Figure 5.41(d) and can also be seen in the un-de-trended signal, Figure 5.41(a). The period of the transit effect is 1.284 days. This candidate is however unlikely to be a planet due to the depth of the signal. There is a strong signal in the difference images at the candidate coordinates, Figure 5.40. The un-phased lightcurve, Figure 5.42, shows in-transit points in four cycles of the transit signal, however the strong nightly trends and other scatter make this less likely to be a transit candidate.

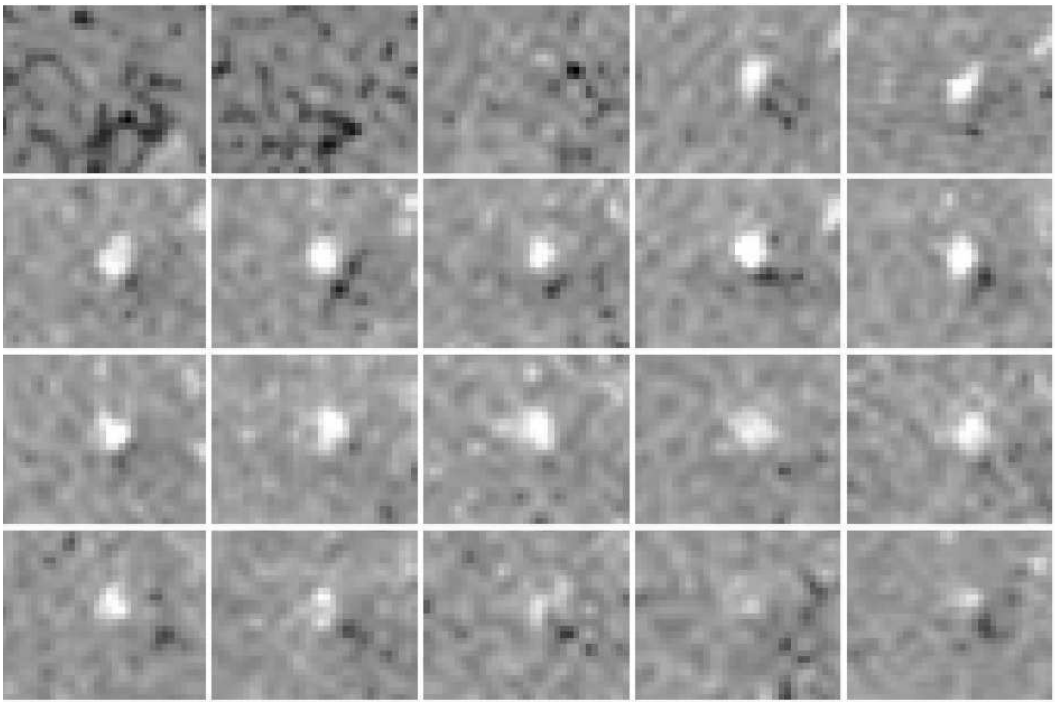
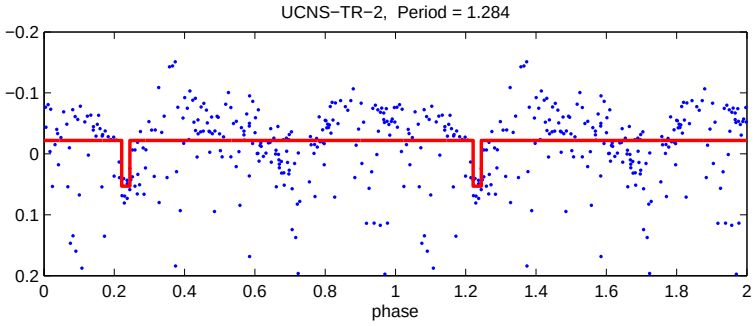
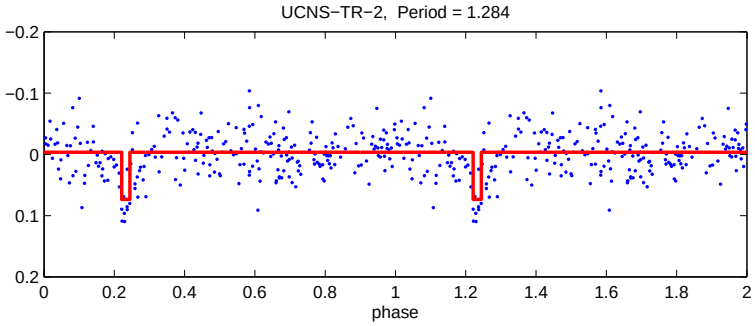


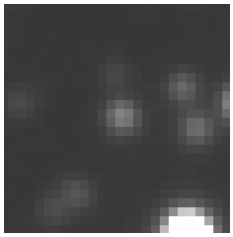
Figure 5.40: **Candidate UCNS-TR-2**: image subtraction sections. The frames show a series of difference images from before (frames 1-3) and through the transit from the re-reduction (note that the pixels here are oversampled by a factor of two). The candidate star coordinates are centred in the frames with each frame being 3.8" by 3.2".



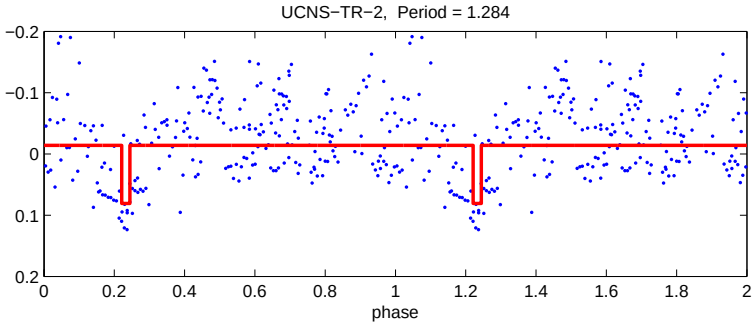
(a)



(b)



(c)



(d)

Figure 5.41: **Candidate UCNS-TR-2**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

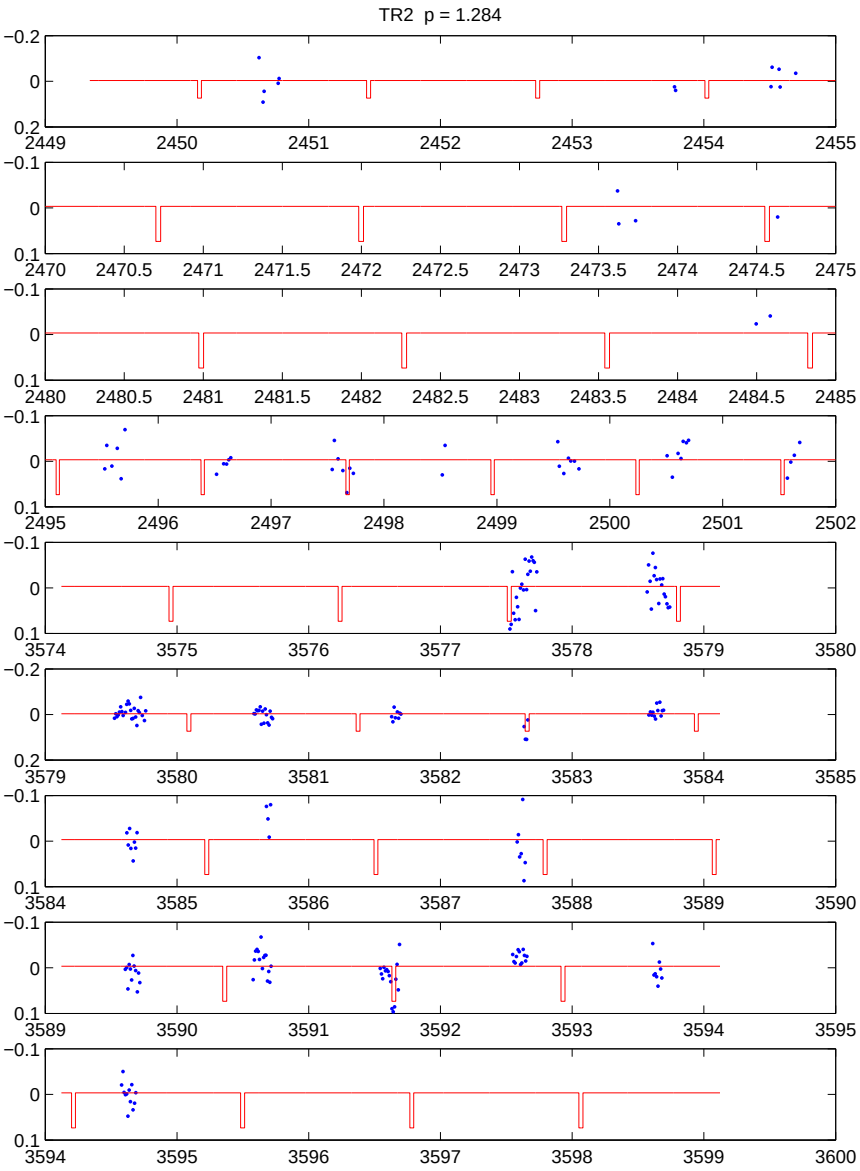


Figure 5.42: **Un-phased candidate UCNS-TR-2** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

5.8.2 UCNS-TR-4

Candidate UCNS-TR-4 has a deep transit signature in the lightcurve (82±9 mmag), Figure 5.44(b), and thus is likely to be something other than a transiting extrasolar planet. The period of the transit signature is 1.9224 days. The same signature has appeared in both the original reduction and the second reduction, Figure 5.44(b,d) and is also present in the noisier un-de-trended original reduction, Figure 5.44(a). The star is faint in the reference frame but is clearly distinguished from neighbouring stars, Figure 5.44(c). Examination of the in-transit difference images revealed a residual in some of the in-transit frames, Figure 5.43. The un-phased lightcurve, Figure 5.45, shows a number of in-transit points through multiple cycles and through the full transit well. This is perhaps the most promising candidate in this research.

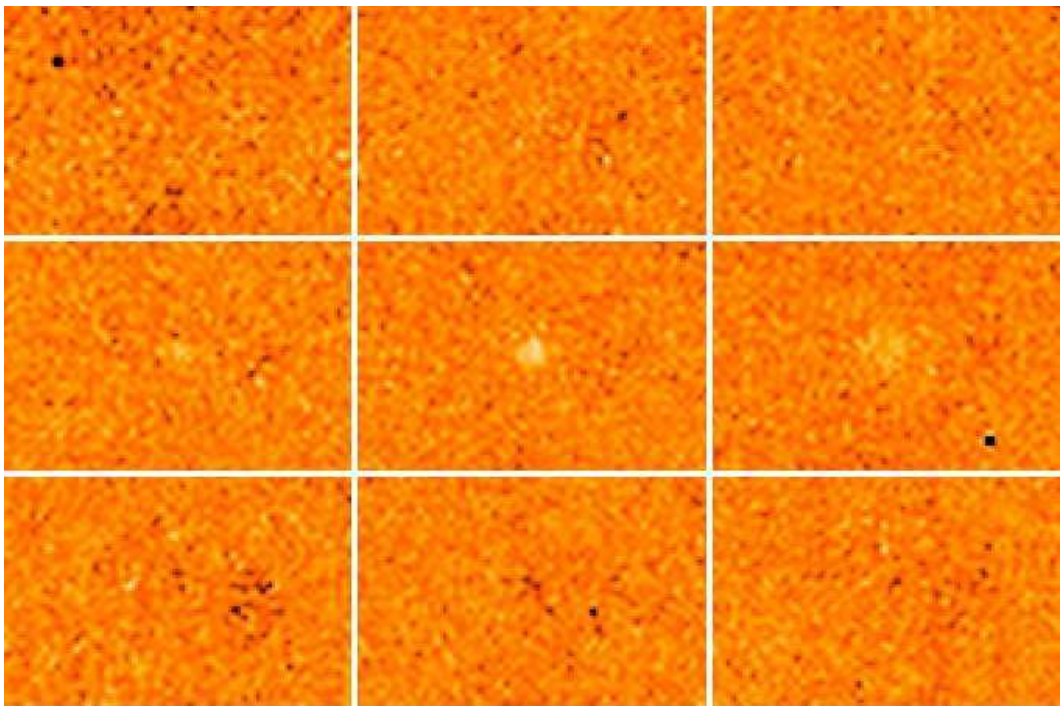


Figure 5.43: **Candidate UCNS-TR-4**: image subtraction sections. The frames show a series of difference images from before (row 1), through one transit (row 2) to after (row 3) from the re-reduction (note that the pixels here are oversampled by a factor of two). The candidate star coordinates are centred in the frames with each frame being 13.1" by 8.7".

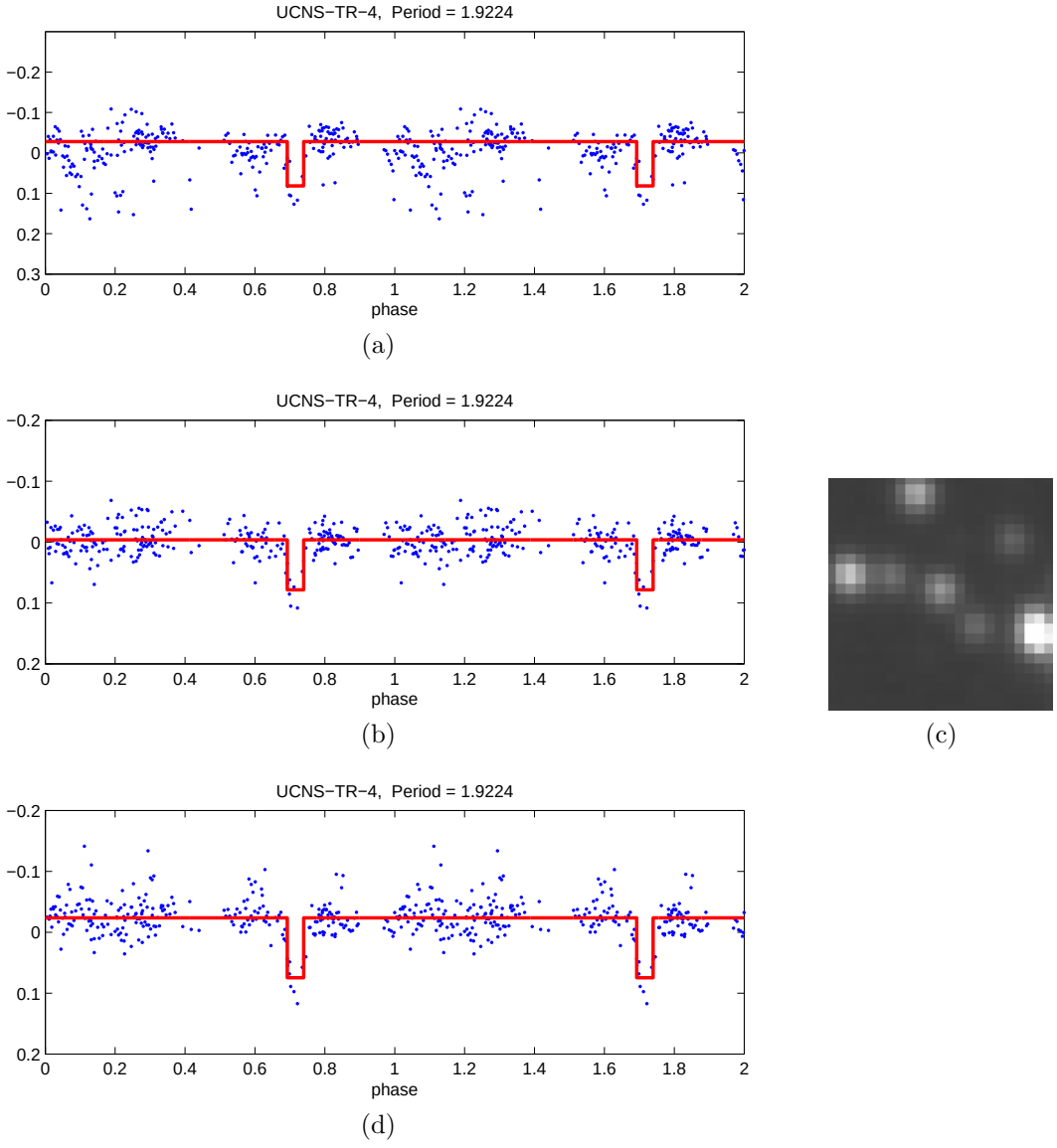


Figure 5.44: **Candidate UCNS-TR-4**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

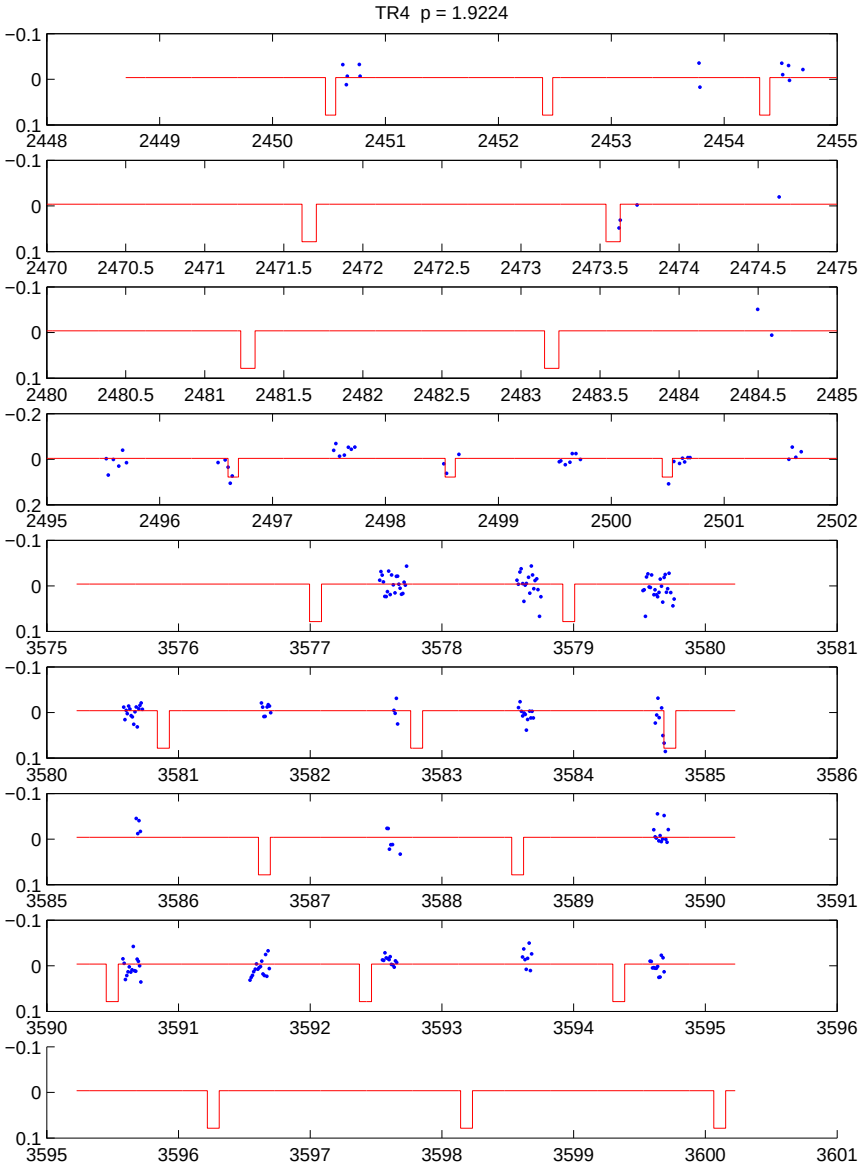


Figure 5.45: **Un-phased candidate UCNS-TR-4** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

5.8.3 UCNS-TR-31

Candidate UCNS-TR-31 seemed a reasonable, though deep (71±8 mmag), candidate from the original lightcurve, Figure 5.46(b). The re-reduction shows the same signal but even deeper (87 mmag), Figure 5.46(d). This signal cannot be clearly seen in the un-de-trended lightcurve, Figure 5.46(a). The transit signal has a period of 2.6269 days with a length of 4.28 hours. The un-phased lightcurve, Figure 5.47, shows large scatter and the in-transit points, while found in multiple cycles, are not conclusive.

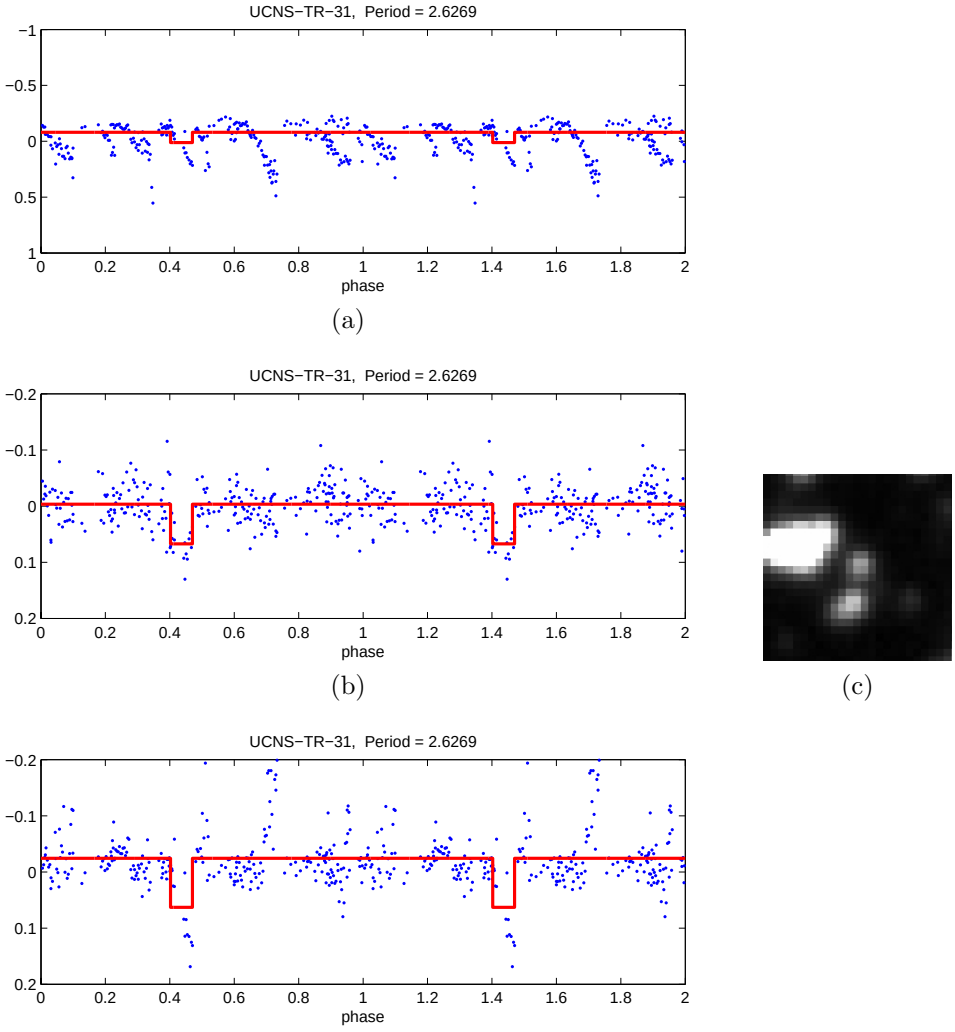


Figure 5.46: **Candidate UCNS-TR-31**: transit candidate lightcurve. Figure (a) shows the un-de-trended original reduction, figure (b) shows the initial de-trended reduction, figure (c) a finder image and figure (d) the lightcurve from the re-reduction without cleaning or de-trending.

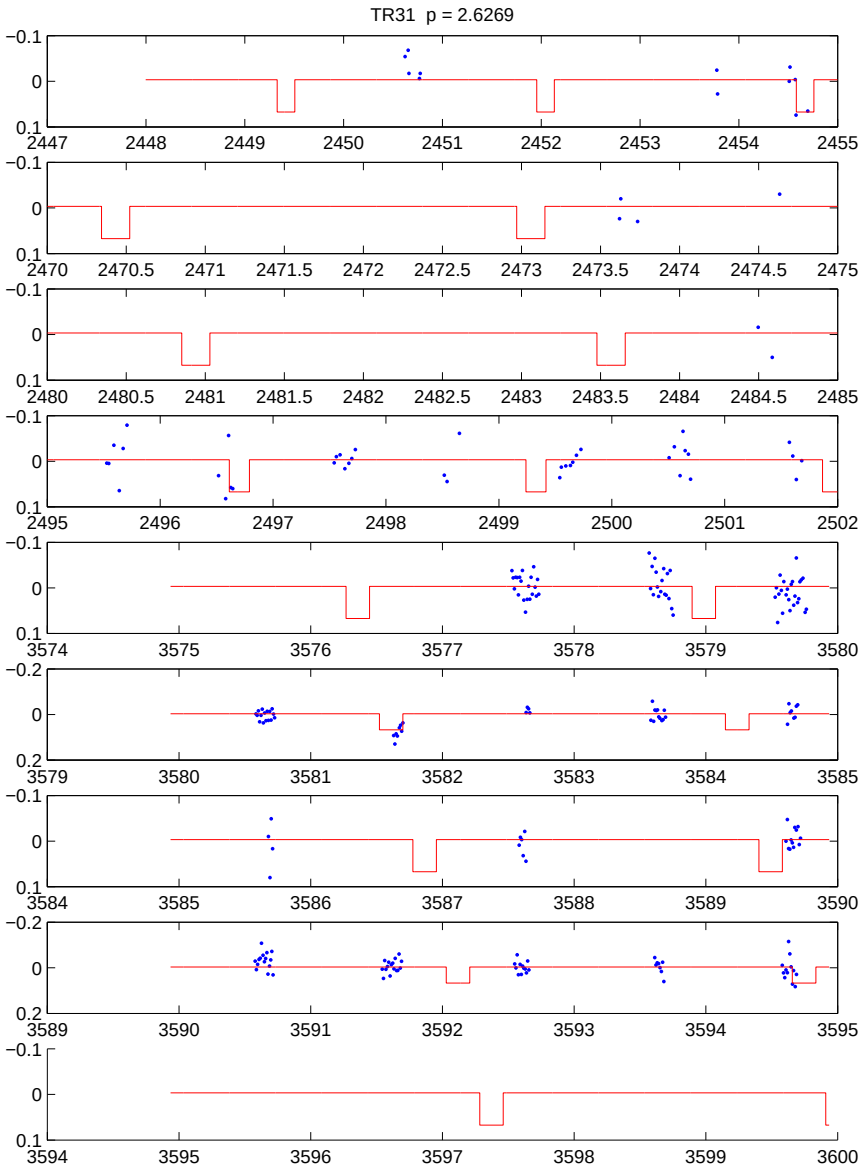


Figure 5.47: **Un-phased candidate UCNS-TR-31** transit candidate lightcurve un-phased and over-plotted with the detected transit signal.

5.9 Summary

From the analysis of the 500,000 plus lightcurves we have found 3 good transit candidates. Candidate UCNS-TR-7 had a transit signal which did not persist in the re-reduction but persists through several cycles in the un-phased data. The transit signal is $\sim$ 25 mmag deep and the star has an R magnitude of 17.69. Candidate UCNS-TR-4 has a transit signal in both the original data and the re-reduction as well as in-transit data points through multiple cycles of the un-phased lightcurve. This candidate is fainter (R = 19.18) and has a deeper transit well (82 mmag) than UCNS-TR-7. Candidate UCNS-TR-13 also seems to be a possible candidate. The un-phased lightcurve only shows in-transit points in one cycle, but it does show a continuous series of points over this cycle. This candidate has a depth of 36 mmag and an R magnitude of 17.27.

In Chapter 2 we discussed the expected results from the proposed transit survey. These predictions indicated a yield of nine or ten planets. These expectations were based on the observations occurring as scheduled. The observations obtained were around a third of that expected and the data points more scattered in time which reduced our detection efficiency. As discussed in §5.4, from simulations we find that the effective number of stars is only 16.6% of the total and thus the expected number of planets reduces likewise to 1.66. The yield from the transit search is as expected.

5.10 Follow-up

In order to confirm the nature of the candidates discussed in §5.9 further observations are needed. In the first instance targeted photometric observations of the individual stars would be best, in particular if this was in multiple filters. Closely spaced continuous observations of several transits would show the shape of the transit and detect blended eclipsing binaries. Such observations would also allow the

use of equations to determine the planet candidate parameters, (Seager and Mallén-Ornelas, 2003) and thus determine whether the lightcurve fits the transit criteria. Multi-colour photometry could also be used to determine whether the candidate is an eclipsing binary as the changes in transit shape through multiple wavelengths is quite distinct from that of eclipsing binaries. Eclipsing binaries generally have large changes in the lightcurve depth between different colours while a transit lightcurve as caused by a planet should not change by more than about 10% in depth.

Further follow-up would involve making measurements of the radial velocity of the candidate stars. The candidates resulting from the transit search have calibrated R magnitudes which are reasonably faint (17.27 - 19.18). However taking radial velocity measurements at medium resolution, about 10 kms^{-1} precision, would eliminate concern that the candidate's signature is produced by a grazing eclipsing binary. The radial velocity signature of a binary is strong enough to be detected by medium resolution spectroscopy. With elimination of false positives via photometric observations, time is more likely to be available on the large instruments required for high resolution spectroscopy of the faint candidates.

Chapter 6

Variable stars

A star is defined as being variable if the observed intensity varies with time. For this photometric study this definition is perfectly suitable but it is noted that in some cases the variability may only be detected in variations in the stellar spectrum.

The photometric variability may be due to the star itself or due to the surrounding environment blocking some of the light from reaching the observer. Intrinsic variables are those whose variability is caused by the nature of the star (i.e. due to stellar pulsations). Extrinsic variables are those with variability caused by the something other than the star, for instance a binary star where one star periodically blocks the light from the other. Variables are in general easier to detect than extrasolar planetary transits due to an often higher amplitude and longer variation of the intensity.

The nature of the observations for this research mean that the detection of short-period variables is highly likely. To this end, in conjunction with the transit search algorithms, several different methods have been employed to detect variable stars, §4.4. Combining the Lomb-Scargle algorithm with the PDM algorithm as a fine tuner for the period gave the best results.

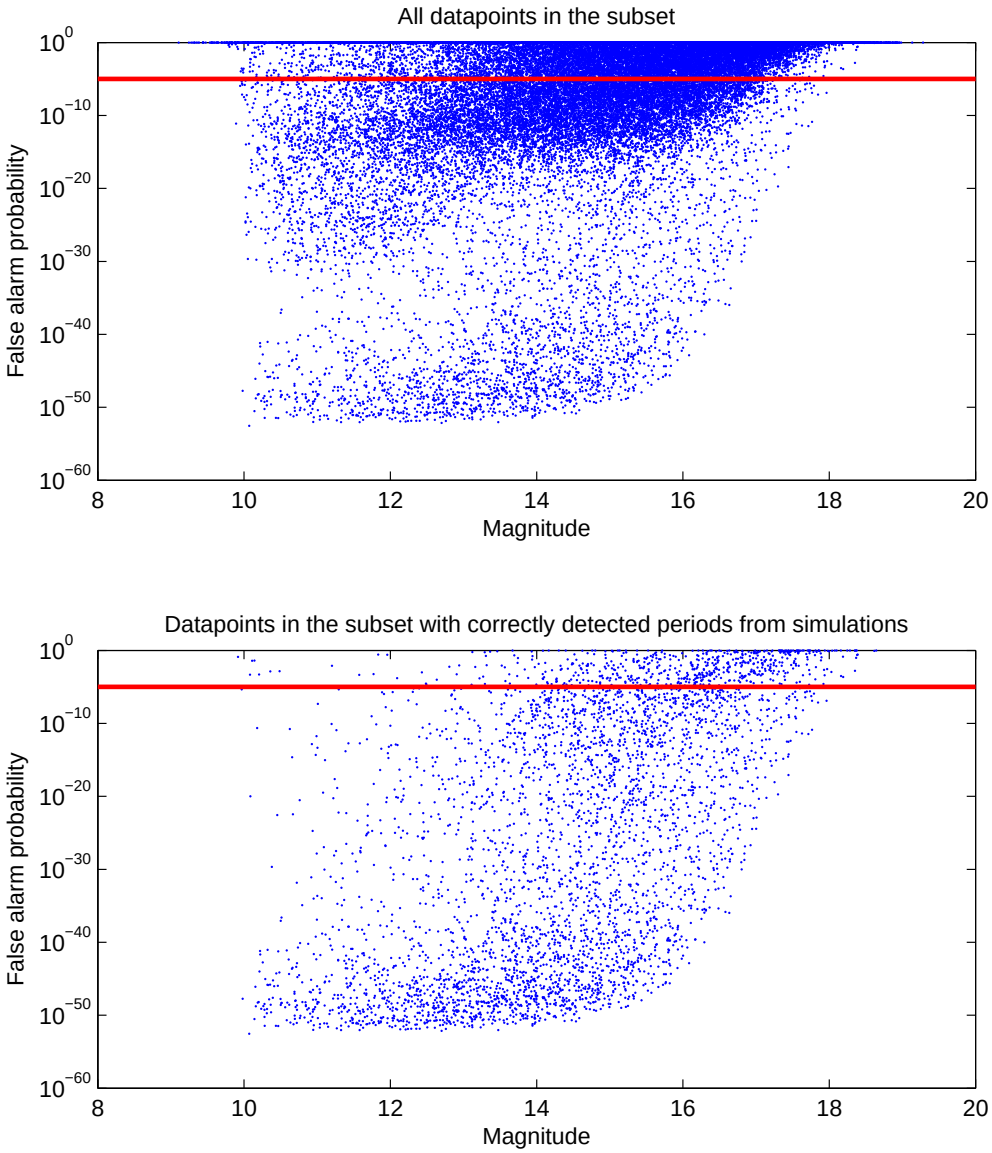


Figure 6.1: **Variable simulations** Lomb-Scargle false alarm probability from insertion of simulated variables into real data. The top panel shows the whole subset, in the bottom panel only those variables detected are plotted.

6.1 Variable simulations

In order to determine the appropriate threshold for variable star detections, simulated sinusoidal variables were inserted into a subset of the real data. The periods range from 0.01 to 10 days with amplitudes varying from 0.05 to 0.5 mags. Figure 6.1 shows the results from using the LSA on the dataset with randomly inserted lightcurves. We find that 80% of these results have false alarm probabilities (FAP) below 10^{-8}. The FAP value decreases as the amplitude increases. The largest number of variables we expect in the dataset will have lesser amplitudes and so it is useful to assume a slightly more pessimistic threshold and so we use 10^{-6}, this will detect a larger number of non-periodic variations but will ensure that most variables are detected. When the seeded data results are compared to the unseeded results we find a large proportion of the unseeded lightcurves with statistics between 0.01 and 10^{-6} therefore using a threshold greater than 10^{-6} is inadvisable.

6.1.1 Variable detection efficiency

From inserting a number (approximately 10,000) of simulated sinusoidal variables into a portion of the real dataset we can determine the detection efficiency of the search algorithm. From Figure 6.2 we see that the efficiency peaks at 90%. Each panel shows the fractional correct detections as a function of magnitude for a different inserted amplitude. Correct detections are defined here as being when the period found is within 10% of the inserted period or an alias and have a FAP below 10^{-6}. In the lower amplitudes we see the detection efficiency drops at the fainter magnitudes though detections are still being made.

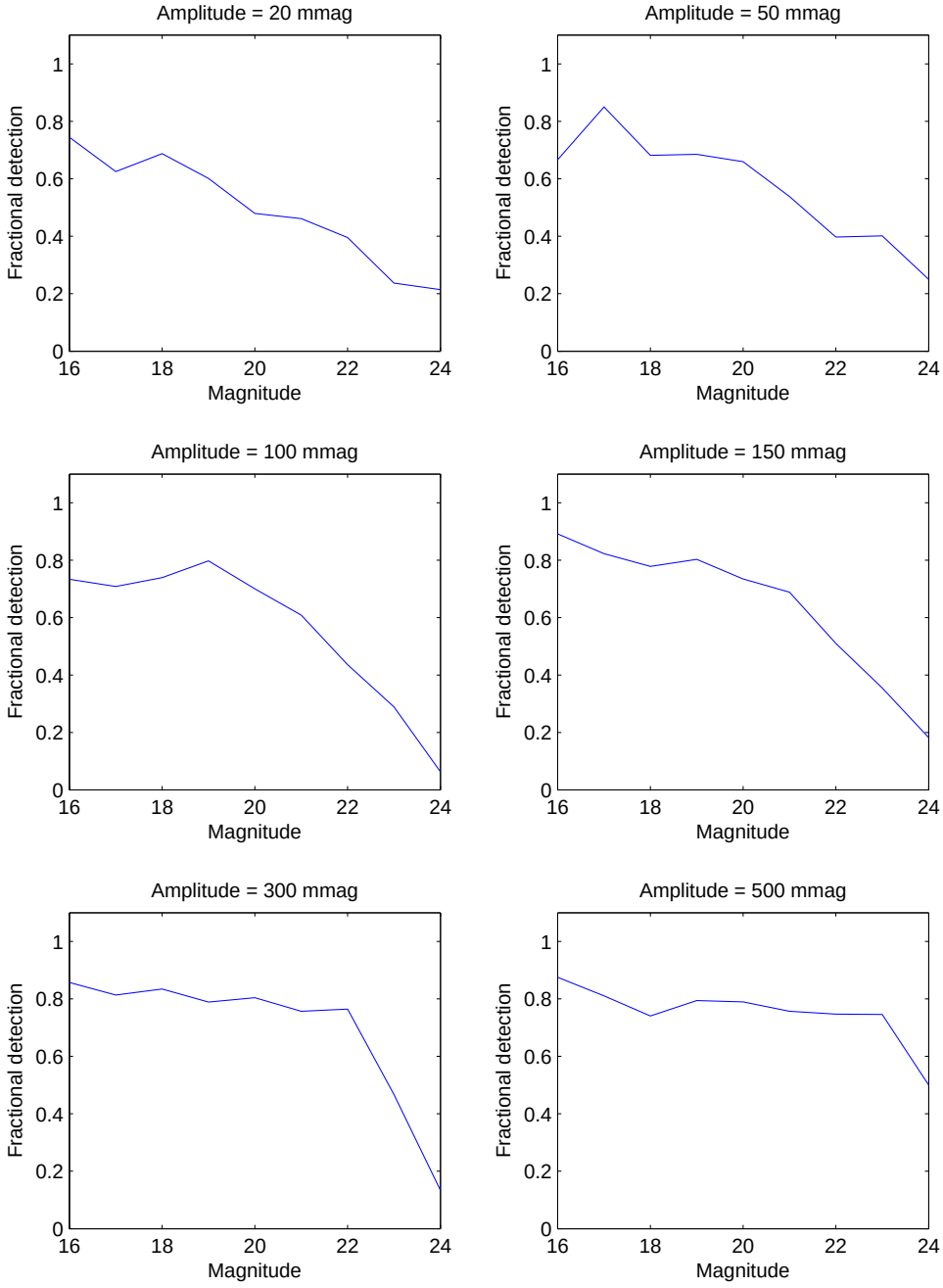


Figure 6.2: **Variable detection efficiency** Results from insertion of simulated sinusoidal variables into real data and searched with the variable detection algorithm - panels show correctly detected variables as a fraction of inserted transits as a function of magnitude for each amplitude.

6.2 Variable search

6.2.1 Lomb-Scargle variable search

The LSA was run on every lightcurve for both the original and detrended datasets (with trial periods ranging on a log scale between 0.01 and 10 days) and the highest power was recorded with the attendant FAP and period. The period range was chosen due to the limits on observations - there are only 16 consecutive days of observations that are well sampled. Longer periods were considered but due to the uneven spacing of the data it is far less likely that a variable with period longer than 10 days (where we will see only a single cycle) will be detected. However as part of the variable search periods are checked for aliases and some longer period variables have been detected in this manner. A log scale was used to decrease the period step size (at periods less than 0.5 days this is at most 0.0012 days) at small periods without slowing the algorithm dramatically. A threshold value is determined below which lightcurves are most likely to be truly periodic. There are some problems with periodic function at 0.5 and 1.0 days where low significances are recorded but no true variations are seen.

Variable candidates are chosen by examining by eye all phased lightcurves with LSA FAP below the threshold 10^{-6}. A large number of non-varying stars may be viewed this way but it is the most accurate method of determining true variables. A few seconds is all that is needed to judge if the lightcurve is truly varying. Simulated variables (§6.1) were used to choose an appropriate threshold value.

Star	Power (detrended)	FAP (detrended)	Period (detrended)	Power	FAP	Period
2310	2.84880911329	1	0.36897759857	2.18898018824	1	0.270395836411
2311	8.82511599954	0.356711932252	0.251188643151	5.90707253319	0.99971818446	0.0732824533139
2312	59.0349564877	6.89779293985e-23	0.993116048421	6.88813952767	0.953206436657	1.47231250243
2313	45.0756467744	7.96455165149e-17	0.997700063823	8.02136431981	0.626786763149	0.0144877185354
2314	6.9058894388	0.950615157308	0.550807696405	5.61142377129	0.999983153054	0.0514043651582
2315	21.0992189374	2.06059023994e-06	1.1220184543	8.52709294339	0.448076778833	3.34965439158
2316	1.42519222626	1	1.27057410521	0.998687831715	1	0.011220184543
2317	54.7811190744	4.85431851344e-21	1.00230523808	10.2702324724	0.0987606516954	0.74644875841
2318	NaN	NaN	0.01	NaN	NaN	0.01
2319	35.3346859141	1.35396327026e-12	0.342767786546	6.60309639559	0.982967045096	0.0162929603264
2320	11.1297376179	0.0430686258513	0.333426412763	7.32119708226	0.862678180747	0.0243220400907
2321	9.66512063771	0.173409375731	0.508159442561	5.89967835333	0.9997347949	0.0990831944893
2322	8.40011383779	0.490754922852	0.990831944893	6.44216362689	0.991639806422	0.0117219536555
2323	NaN	NaN	0.01	NaN	NaN	0.01
2324	69.7093196252	1.59548654098e-27	0.997700063823	10.2430014288	0.101344004794	0.390840895792
2325	6.13212818107	0.998533859493	0.143879857826	6.36600877152	0.994276515873	0.0770903469064
2326	9.73285109705	0.163035399744	0.14256075936	7.18720175589	0.896710678993	0.0106905487922
2327	22.6834544497	4.22636399072e-07	0.997700063823	6.05739313964	0.999116885595	0.129419584145
2328	27.2023359114	4.60724825495e-09	1.00230523808	8.70656753611	0.391459985181	0.157398286447
2329	46.7413623417	1.50574092294e-17	0.995405417352	7.00747723189	0.933953889792	0.0779830110523
2330	23.8157931282	1.36206695827e-07	1.11173172728	6.21436034517	0.997544684947	0.135518941235
2331	9.85509180216	0.145718491282	1.02801629813	5.39670668987	0.999998796018	0.872971368388
2332	14.4853706827	0.00153467382949	0.464515275223	8.32357286616	0.517373584148	0.112719745618
2333	30.7565736521	1.31781903421e-10	1.01859138805	5.80601311074	0.999882001551	0.0479733448637
2334	3.91840240801	1	5.23600436586	2.62491250896	1	2.9444216338
2335	23.0431434414	2.9495512939e-07	1.09647819614	6.35480142613	0.994600339439	0.0157398286447

Table 6.1: **Lomb-Scargle results matrix** Sample of results as shown in the matrix resulting from the Lomb-Scargle algorithm on a subset of lightcurves.

6.2.2 PDM method

In the PDM method (§4.4.3) a minimum value is desired, Figure 6.3 shows the periodogram for a sample artificial lightcurve. The minimum centered at 4.01 days indicates the best period for this lightcurve.

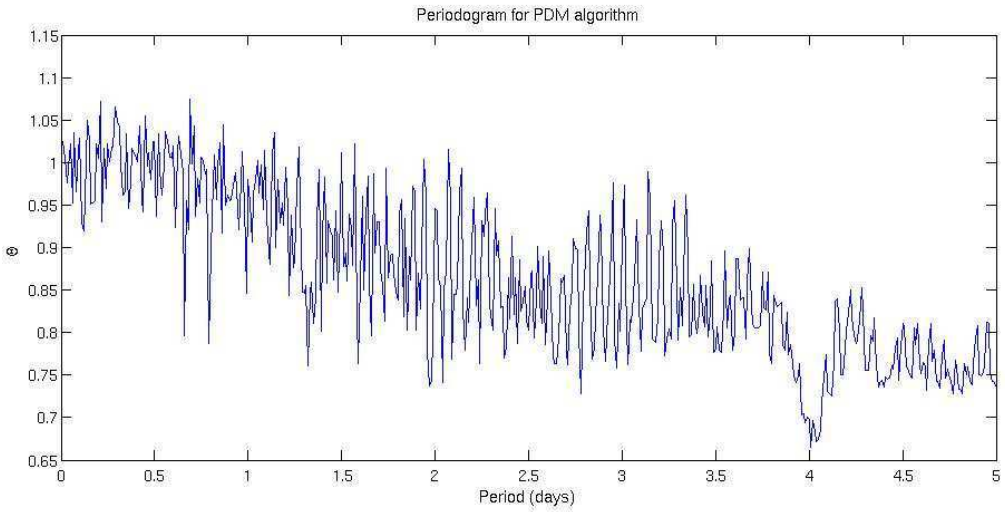


Figure 6.3: **PDM results** Resultant periodogram from a single star through the PDM algorithm. Shows the test statistic against the period.

The output from the original PDM code created a periodogram as in Figure 6.3 with a small table containing the results. The code used in the analysis of all the lightcurves in the data set instead produces a matrix containing the resulting best parameter set as well as the significance and $\Delta\chi^2$ found for that set. Using these, the lightcurves most likely to contain a candidate periodic signal can be identified. A sample of this matrix can be seen in Table 6.2.

Lightcurve	Resultant Period	Significance	$\Delta\chi^2$
6400	0	0	0
6401	0	0	0
6402	2.24388	111.898	0.136774
6403	4.10204	0.352038	0.353675
6404	2.50035	0.0468217	0.467538
6405	7.17794	2.41582	0.0930848
6406	1.11429	1.0982	0.00114487
6407	4.10204	7576.63	0.0498855
6408	2.09894	0.0357131	0.297203
6409	1.57036	5.72294	0.457753
6410	1.80717	0.0103024	36.7822
6411	1.00462	0.111722	2.41767
6412	1.25893	0.0128646	1.27345
6413	2.37137	0.00808558	10.0304
6414	1.98153	11423.2	1890.09
6415	1.90108	0.0339957	37.0689
6416	2.76058	0.0201485	2.41136
6417	3.63915	0.00876728	471.648
6418	6.95024	0.109781	0.0173166
6419	2.09411	5.441e+06	15.0887
6420	1.45546	0.00646217	0.75431
6421	0	0	0
6422	1.06170	0.00946305	0.199423
6423	1.90985	0.179947	0.310265
6424	2.02302	0.0977844	0.96378
6425	0	0	0
6426	2.37137	135.868	9.89215
6427	0	0	0
6428	4.88652	0.0226514	0.182914
6429	2.05589	13.0433	1.93598
6430	0	0	0

Table 6.2: **PDM results matrix** Sample of results as shown in the matrix resulting from the PDM algorithm on a subset of lightcurves.

6.2.3 Refining periods

Given a large range of periods we searched every lightcurve using trial periods on a log scale from 0.01 to 10 days with a step size 0.025% of the trial period (0.000025 at 0.01 days to 0.025 at 10 days). Periodic signals are detected but in a large number of cases the true period is not defined accurately. In order to obtain a more accurate period the algorithm is run a second time on those lightcurves with FAP below the threshold. The period is restricted to 0.2 on either side of the originally determined period and the step size is decreased to 0.00001 day. For this research we ran the LSA first and then two fine searches, the PDM and LSA again. The LSA picks up both sinusoidal and non-sinusoidal periodic functions but the PDM method is better at determining the correct period for non-sinusiodal oscillations. Following the fine search we visually examined the phased lightcurves and judged whether each candidate star is truly a periodic variable.

All lightcurves from the mosaic field observations in 2002 and 2005 were analysed as above. Contained in the folowing sections are the raw results from these algorithms. The algorithms have resulted in a large number of variable star identifications. These are discussed in §6.3 and §6.4 with lightcurves in Appendix B. More than 1,000 new variable stars have been identified and catalogued.

6.3 Variable catalogue

A total of 1475 variable lightcurves have been found only one of which had previously been known (see §6.3.2 for more details). The catalogue includes a large number of eclipsing binaries with 143 of Agol type, 336 of β Lyrae type and 588 contact eclipsing binaries (W Uma type). Also found are 9 possible Cepheids (DCEP) and a large number of un-categorized pulsating stars (PUL). The catalogue supplies a large

amount of data on variable stars in the Galactic plane. Unusual and unexpected results can be used to test theories and concentrations of variable stars.

The catalogue contains the periods found and then refined by the LSA and PDM algorithm. In the variable identification process the phased lightcurves resulting from searching the detrended data as well as the original data were examined. In almost all cases the variation was found in the un-detrended data and the period found from this was then recorded. The lightcurves were individually examined and assigned a type. The periods were also checked for aliases - in a large number of cases where the star is an eclipsing binary the periods needed multiplying by 2 to obtain the correct period. The catalogue contains columns of star identification, right ascension, declination, calibrated magnitudes of R, B-R and R-I where available, the closest period, type and any notes. The variable stars have been given identifications UCNS-V-# (University of Canterbury Norma Search Variable). The acronyms used are in common usage and are defined by Table 6.3. A classification that is not certain is followed by a colon (i.e. EA:), a star of more than one type is given both designations (i.e. EA + PUL) and where there is ambiguity in the class the two types are combined with a slash (i.e. EA/PUL). The notes column contains any other information gleaned about the star. The complete table is found in Appendix A but similar tables have been used further in this chapter to highlight particular features. Table 6.4 gives the first 20 lines of the catalogue as a sample.

Acronym	Star type
EA	eclipsing binary, Agol type
EB	eclipsing binary, β Lyrae type
EW	contact eclipsing binary, W Uma type
CV	Cataclysmic variable
DCEP	Cepheid type
PUL	pulsating types, including δ Scuti stars

Table 6.3: **Variable star acronyms** Table defining the variable star designations used in the variable star catalogue.

ID	Subset	RA	DEC	Calibrated Magnitudes			Period	Type	Notes
				R	B-R	R-I	(days)		
UCNS-V-1	1A	16:20:37.72	-53:25:26.5011	17.84	1.81	1.90	2.18480	EB	
UCNS-V-2	1A	16:20:52.98	-53:25:24.0178	19.68	1.96	1.71	0.43817	EW	
UCNS-V-3	1A	16:20:41.89	-53:25:07.5671	20.27	2.30	2.23	0.67577	EW	
UCNS-V-4	1A	16:20:47.06	-53:25:06.8219	18.66	1.67	2.05	0.41308	EW	2 max
UCNS-V-6	1A	16:20:49.68	-53:24:54.9648	15.97	1.26	1.64	8.45680	PUL:	
UCNS-V-7	1A	16:20:41.60	-53:24:52.8653	17.94	1.58	1.76	1.87580	EB	
UCNS-V-9	1A	16:20:50.50	-53:24:40.1987	19.05	1.33	2.28	0.51056	EW	flat bottom
UCNS-V-10	1A	16:20:50.55	-53:24:38.6527	20.27	2.05	2.06	0.51090	CV/PUL	
UCNS-V-11	1A	16:20:46.81	-53:24:35.0235	19.31	2.16	1.82	0.40644	EW	
UCNS-V-13	1A	16:20:52.65	-53:24:27.0178	18.64	2.19	2.17	0.86336	EA	
UCNS-V-14	1A	16:20:40.27	-53:24:26.6044	18.74	1.87	1.99	1.46712	EB	
UCNS-V-15	1A	16:20:51.95	-53:24:24.8163	18.40	1.78	1.69	0.36636	EW	
UCNS-V-16	1A	16:20:51.97	-53:24:24.0444	18.02	1.70	1.86	0.36654	EW	
UCNS-V-17	1A	16:20:37.33	-53:24:18.8611	17.92	2.62	2.21	2.18690	EA:	
UCNS-V-18	1A	16:20:51.02	-53:24:11.9551	18.48	2.33	2.28	6.45360	PUL:	
UCNS-V-19	1A	16:20:34.96	-53:24:11.0538	19.15	1.88	1.87	2.19202	EA	
UCNS-V-20	1A	16:20:35.35	-53:24:10.5681	19.26	1.80	1.90	0.39752	EW	
UCNS-V-21	1A	16:20:36.55	-53:23:58.2339	17.98	2.24	2.11	4.83100	EB:	
UCNS-V-22	1A	16:20:35.57	-53:23:45.1797	16.00	1.81	1.80	2.13320	EB:	
UCNS-V-23	1A	16:20:54.36	-53:23:34.1039	18.90	1.85	1.85	1.12340	CV/PUL	
UCNS-V-24	1A	16:20:43.49	-53:23:31.1435	20.47	2.74	2.76	2.18060	EB	

Table 6.4: **Variable star catalogue (sample)** Catalogue of variable stars detected in the dataset. Tabulated are star ID, coordinates, magnitudes, period and star type.

6.3.1 Variable stars on the colour-magnitude diagram

Many variable stars can only be classified from photometric measurements by examining their position on a colour-magnitude diagram. In our data we have a main sequence on the colour-magnitude diagram as well as a number of other field stars. In the following figures (6.4 to 6.11) we have overplotted the positions of identified variable stars on the colour-magnitude diagrams. Those variable stars which do not lie on the main sequence are discussed in §6.4.2.

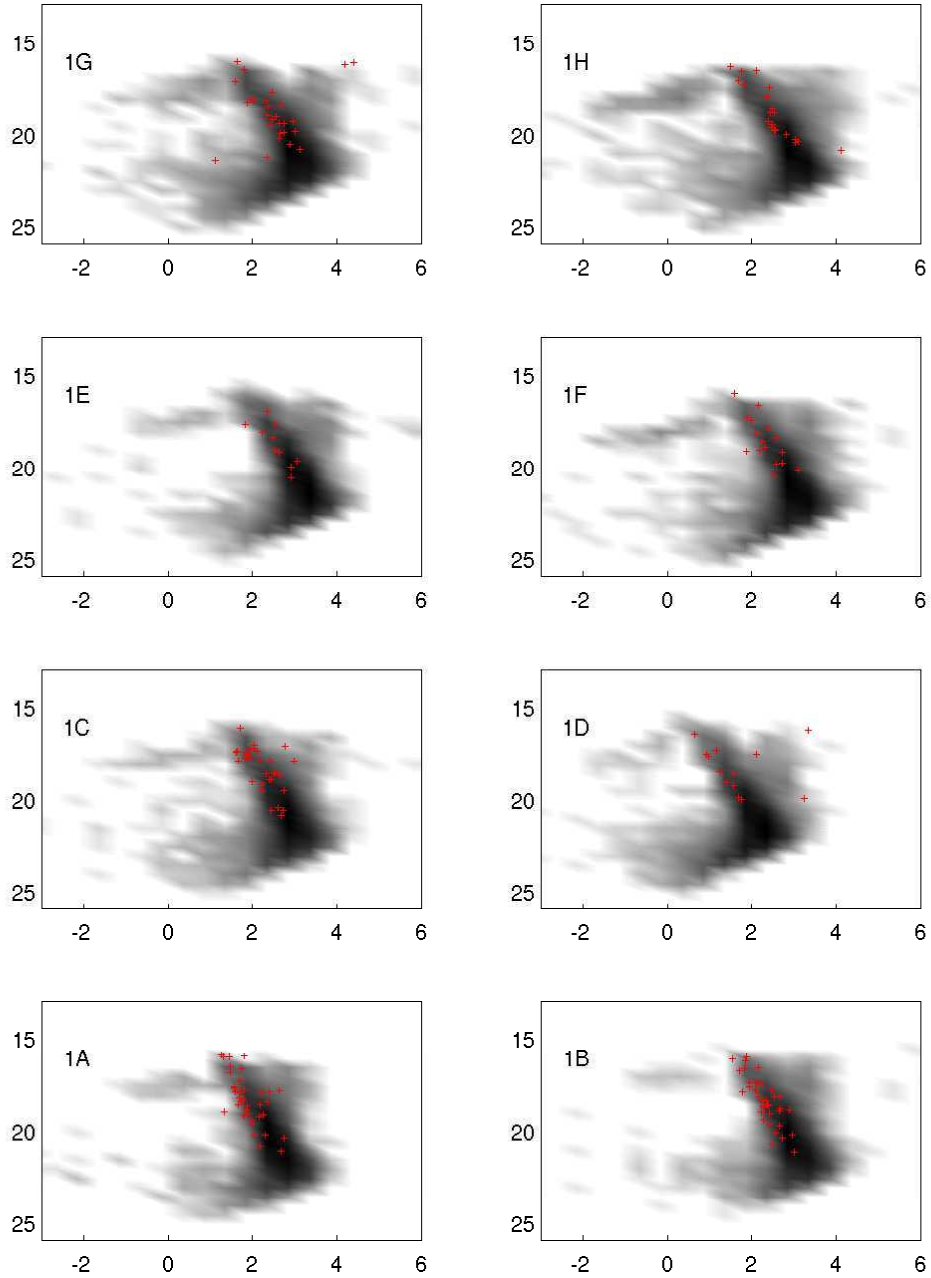


Figure 6.4: **Variable stars on the colour-magnitude diagram (subset 1)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

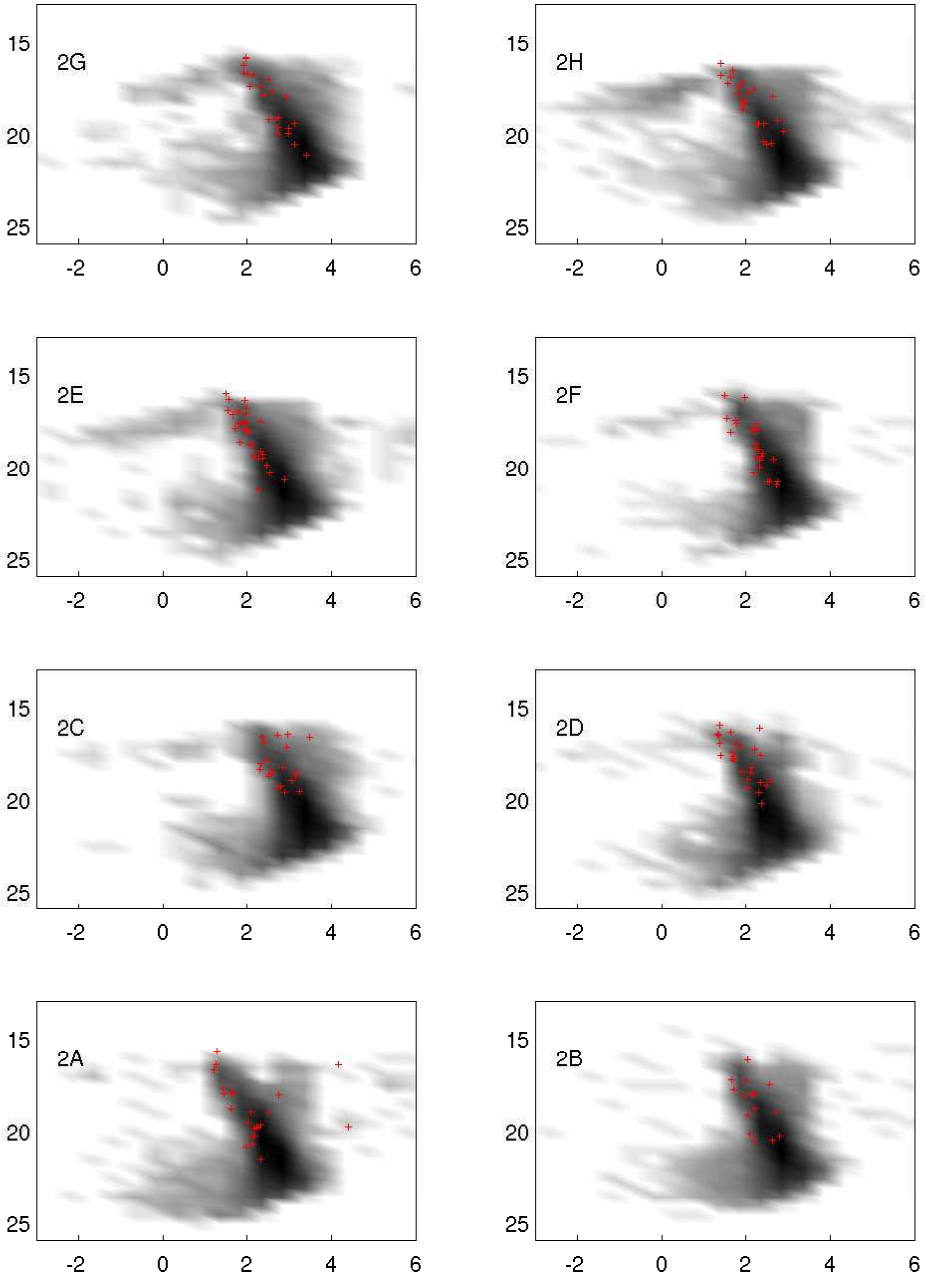


Figure 6.5: **Variable stars on the colour-magnitude diagram (subset 2)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

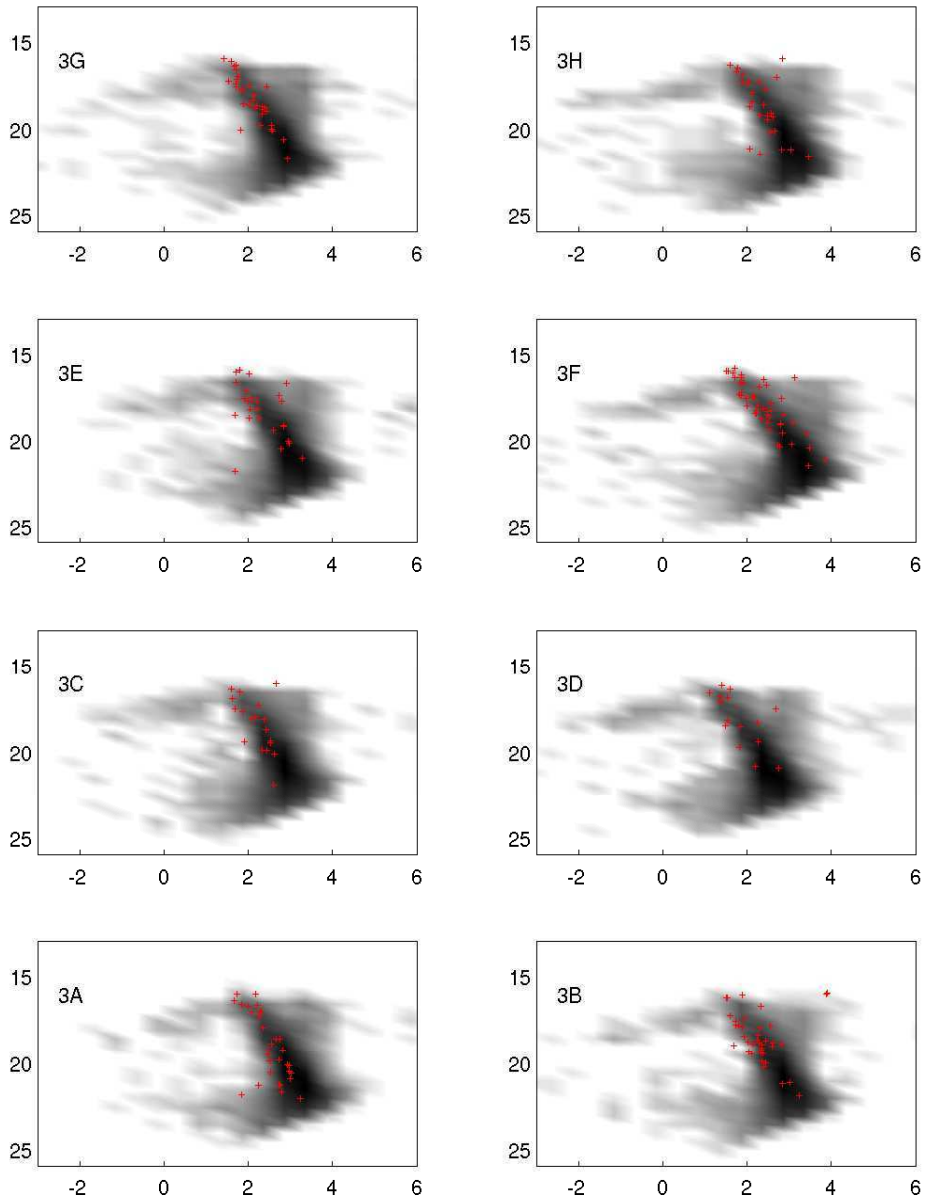


Figure 6.6: **Variable stars on the colour-magnitude diagram (subset 3)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

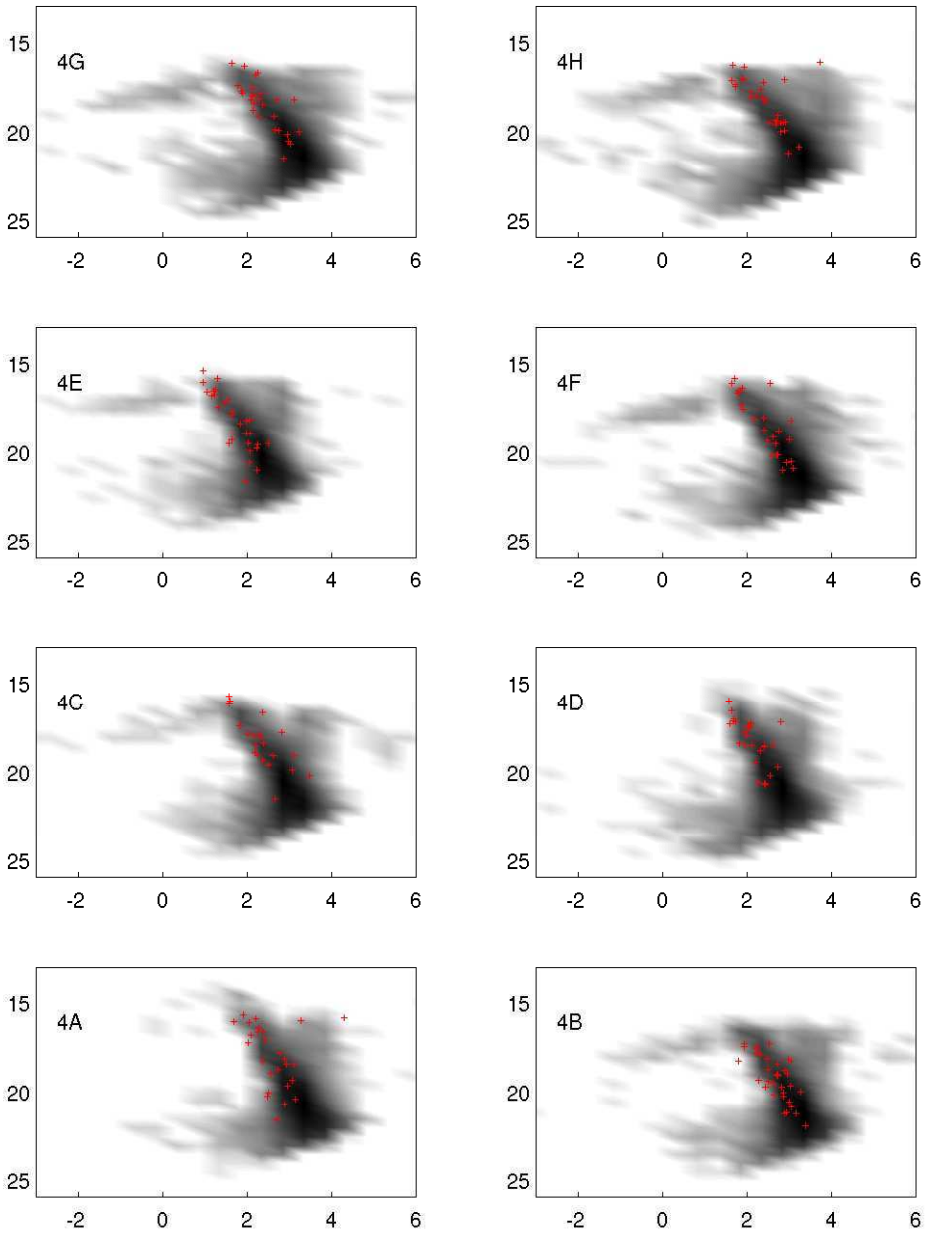


Figure 6.7: **Variable stars on the colour-magnitude diagram (subset 4)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

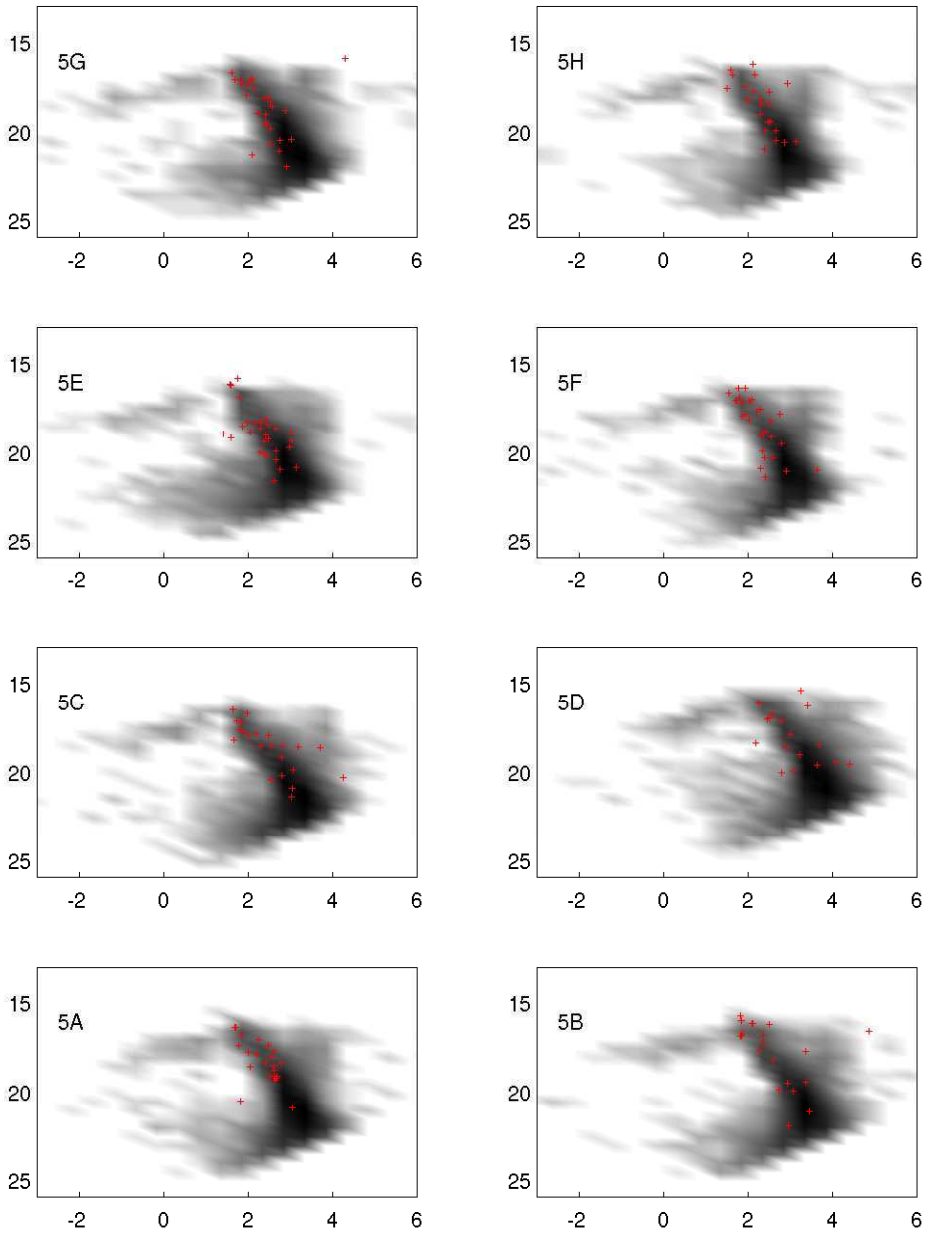


Figure 6.8: **Variable stars on the colour-magnitude diagram (subset 5)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

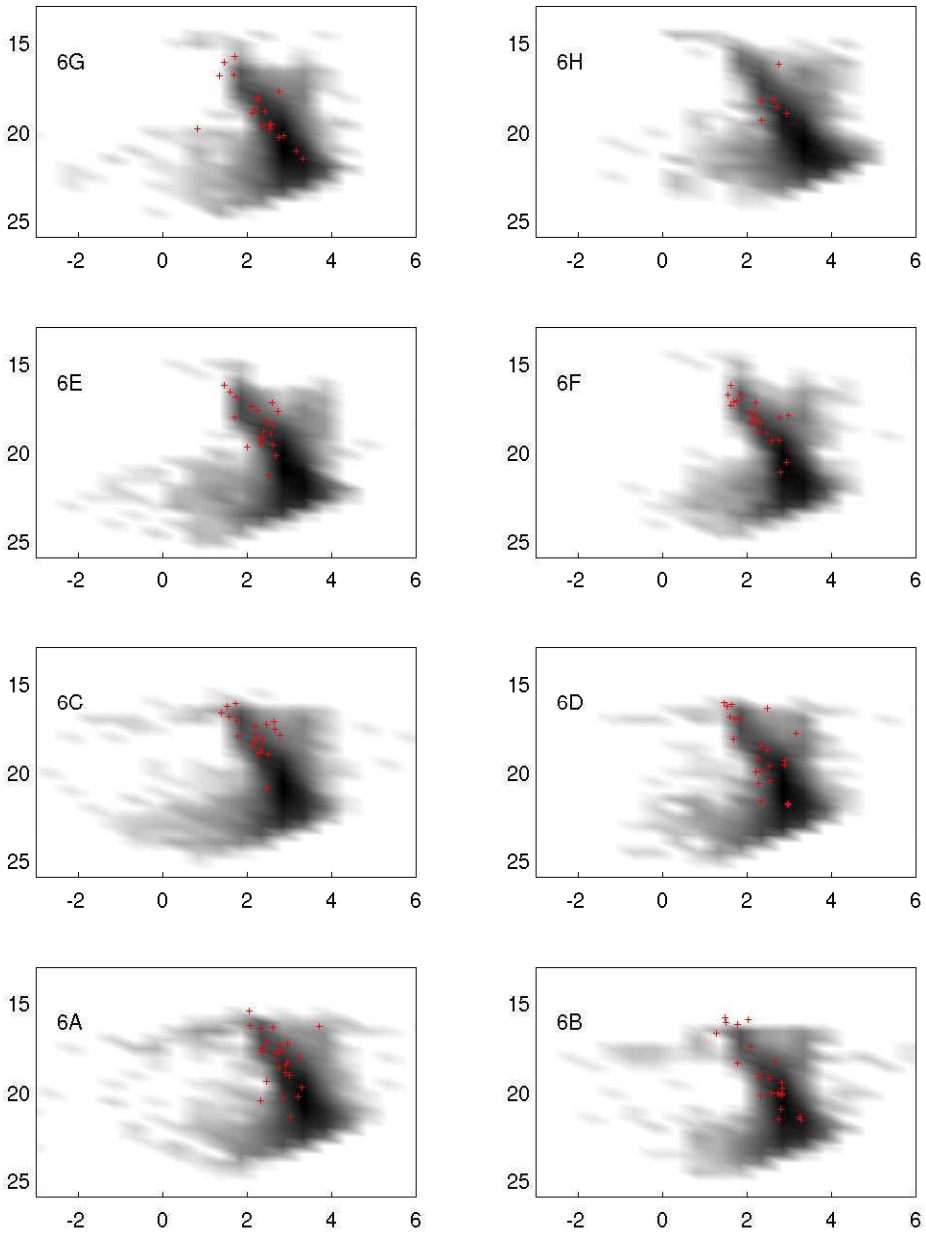


Figure 6.9: **Variable stars on the colour-magnitude diagram (subset 6)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

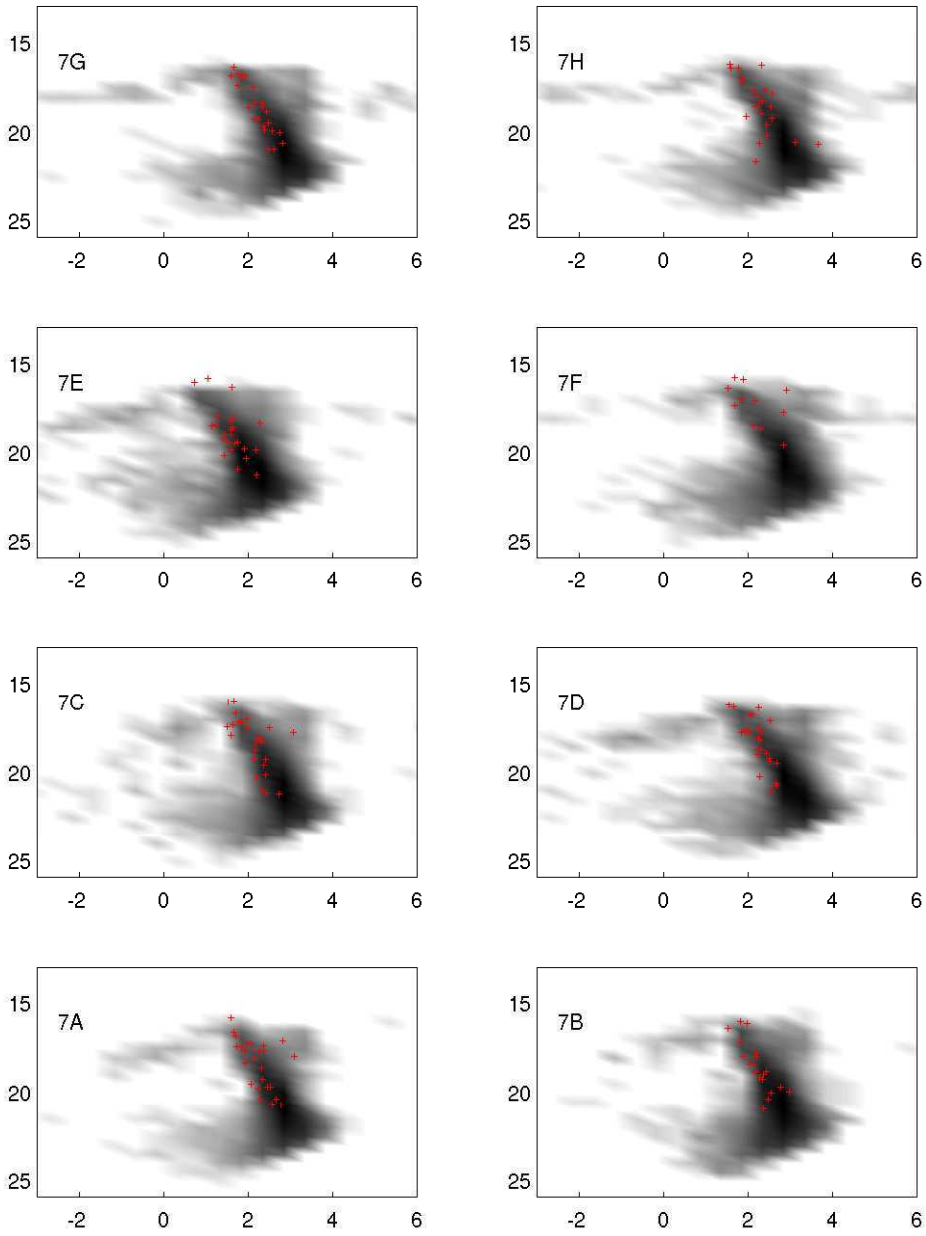


Figure 6.10: **Variable stars on the colour-magnitude diagram (subset 7)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

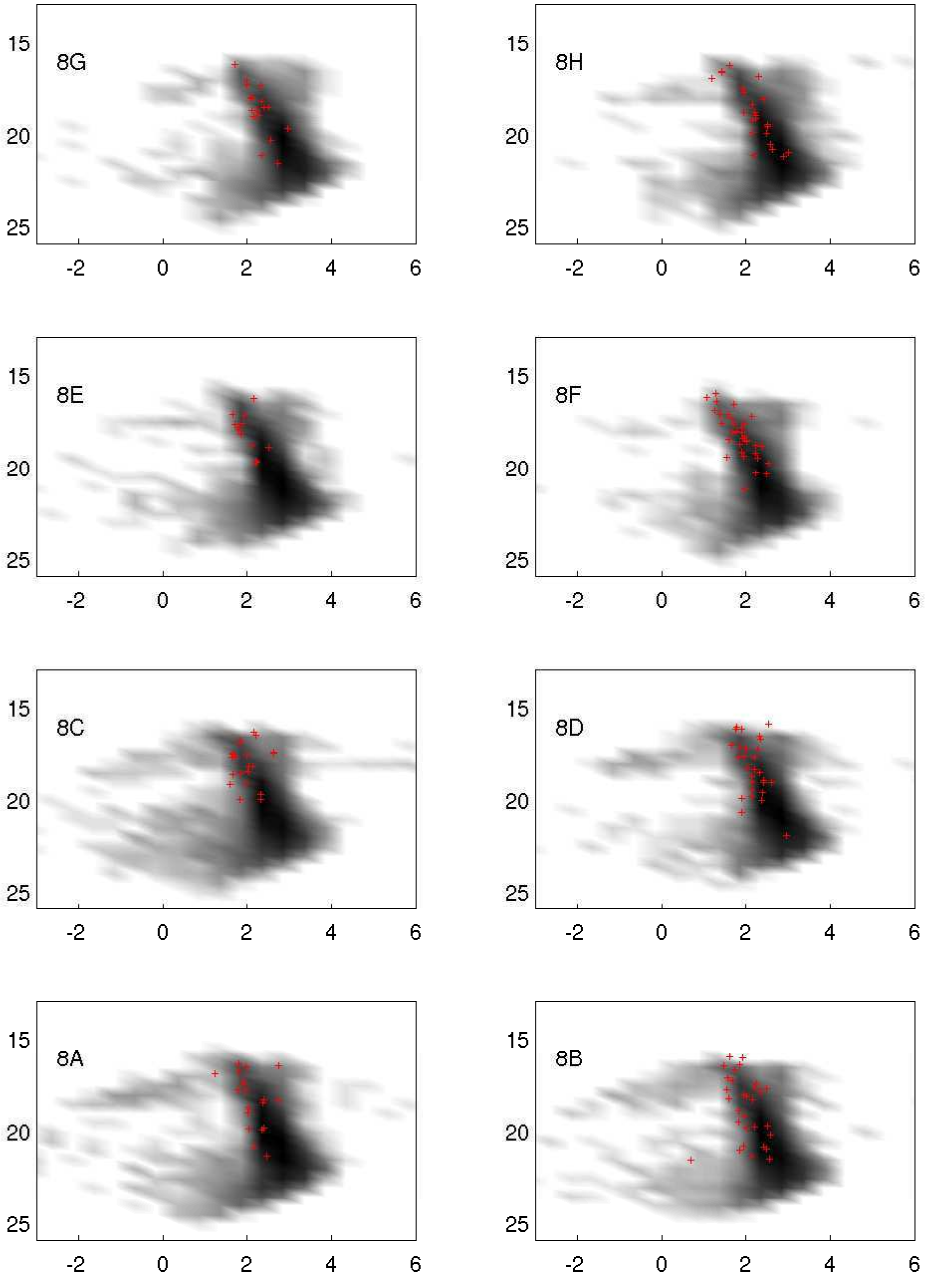


Figure 6.11: **Variable stars on the colour-magnitude diagram (subset 8)** Each panel shows the colour-magnitude diagram (B-R vs R) for the stars of a sub-image with calibrated magnitudes. Each colour-magnitude diagram has the variable stars from that sub-image overplotted with +.

6.3.2 Previously known variables

After the variable stars were identified and the catalogue compiled the x,y pixel coordinates were mapped to sky coordinates and existing variable star catalogues were used to cross-correlate our catalogue to determine previously detected variable stars. Using the *VizieR* (Ochsenbein et al., 2000) on-line database we queried the Combined General Catalogue of Variable Stars (GCVS) (Samus et al., 2004) which includes those stars labelled New Suspected Variable (NSV) and also the All Sky Automated Survey (ASAS) (Pojmanski and Maciejewski, 2004) which has a catalogue of Southern Variables. A 30' cone search was performed at the pointing coordinates.

The search returned thirteen possibilities and so these were examined individually by comparing the reference images to the DSS image of the region displayed by Aladin (Bonnarel et al., 2000) overlain with the variable stars found by *VizieR*. Four of the thirteen fell outside the frame of the observations and three others appeared between the CCD chips. One set of coordinates belonged to an ASAS V band lightcurve and was not detected in our images. The remaining five were detected in our images but of the five, four were strongly saturated and therefore not detected by the DAOPHOT reduction (see Chapter 3) and the final star was saturated in most frames and the lightcurve contained only a few points.

The detected star was HO Nor, an Agol type eclipsing binary with a period of 2.12317 days. The initial variable search (LSA) was performed on all the stars before the fine search was performed solely on those with a sensible number of datapoints and false detection probability below the threshold. HO Nor was not detected in the search due to the lack of points but has since been included in the catalogue (UCNS-V-1452). The lightcurve corresponding to the coordinates of HO Nor is shown in Figure 6.12. Despite the small number of points, it does appear that one of the points obtained is in the primary eclipse of the binary. The period found by the coarse Lomb-Scargle search is close to the known period.

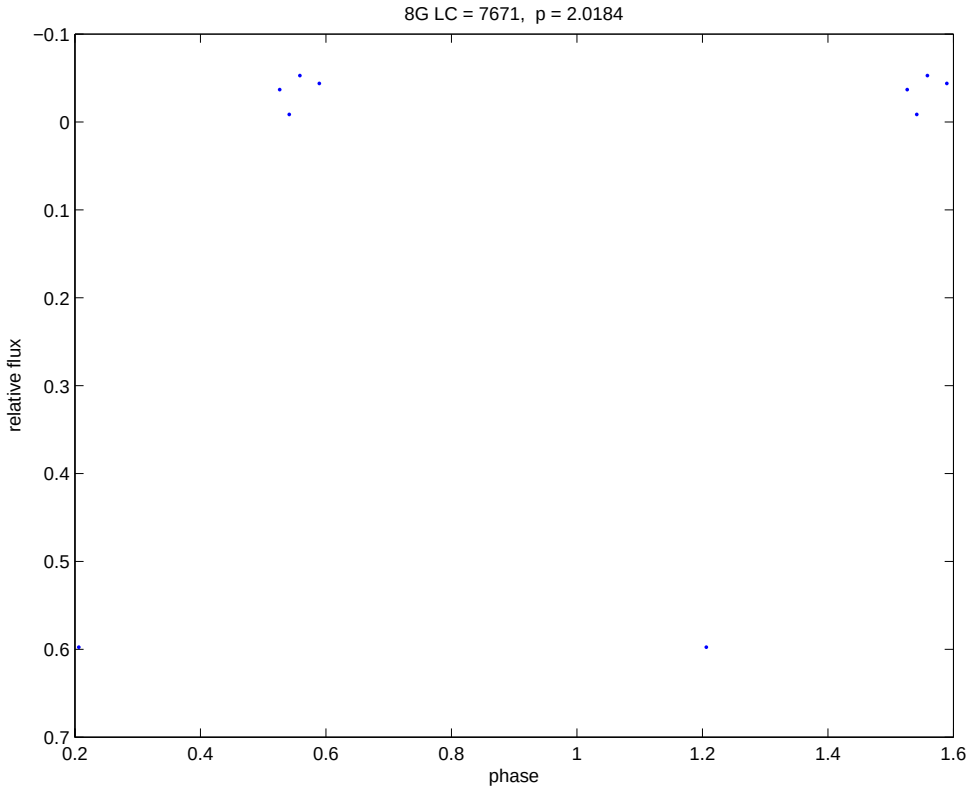


Figure 6.12: **Lightcurve of known variable HO Nor** Phased lightcurve (with found period) of HO Nor with the points found in unsaturated images. The true period of HO Nor is 2.12317 days.

The lack of previously known variables means that almost 100% of the variables found from this survey are unpublished and previously unknown.

6.4 Discussion

In this section we discuss those variable stars which have more unusual properties. Several of the variable stars, in particular the short period contact binaries, appear in more than one subsection.

6.4.1 Contact eclipsing binaries with unusual properties

Contact eclipsing binaries (EW) are amongst the most common variable stars. Contact binaries are detected at a rate of approximately 1 in 500 FGK dwarfs (Rucinski, 2006), and this is born out in this catalogue with 588 EW type stars out of 1475 variable stars (with an effective total number of stars searched for variable stars of approximately 360000 (see §6.1.1, 6.4.3)).

Due to the similarities in lightcurve morphology of BY Draconis type stars and contact eclipsing binaries, without further information we are unable to determine if any of the stars categorized as EW are in fact BY Draconis stars. There are a number of EW stars with lightcurves indicating the possibility of starspots.

In the list of EW type stars there are a number with maxima at different brightness, this is possibly indicative of semi-detached star systems and so these stars may have been mis-categorized. These variables are listed in Table 6.5 with lightcurves in Figures 6.13 and 6.14.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-4	1A	16:20:47.06	-53:25:06.8219	18.66	1.67	2.05	0.4131	EW
UCNS-V-81	1C	16:20:34.42	-53:20:30.1431	17.77	1.95	0.68	0.3164	EW
UCNS-V-220	2B	16:19:26.09	-53:23:07.7984	17.35	2.00	0.43	0.3629	EW
UCNS-V-240	2C	16:19:39.44	-53:19:3.5470	18.12	2.34	0.59	0.3500	EW
UCNS-V-333	2G	16:19:37.64	-53:11:39.8602	18.06	2.94	0.43	0.3213	EW
UCNS-V-369	3A	16:18:59.77	-53:25:09.8420	16.74	2.00	0.19	0.2779	EW
UCNS-V-472	3E	16:19:00.24	-53:16:15.5261	17.58	2.08	0.61	0.3683	EW
UCNS-V-493	3F	16:18:36.50	-53:17:09.0202	16.23	1.87	1.75	0.2725	EW
UCNS-V-528	3F	16:18:22.90	-53:14:53.5032	17.55	1.97	1.72	0.2577	EW
UCNS-V-536	3G	16:18:40.88	-53:13:35.2199	16.68	1.69	0.45	0.2602	EW
UCNS-V-540	3G	16:18:50.54	-53:13:15.8250	18.10	2.13	0.52	0.3485	EW
UCNS-V-558	3G	16:18:54.13	-53:11:08.7186	18.95	2.33	0.85	0.2487	EW
UCNS-V-570	3H	16:18:17.36	-53:12:29.4420	21.25	2.81	1.07	0.3532	EW
UCNS-V-572	3H	16:18:17.49	-53:12:27.6664	17.42	2.01	0.83	0.2998	EW
UCNS-V-604	4A	16:18:00.36	-53:24:10.6221	17.14	2.45	0.45	0.3846	EW
UCNS-V-608	4A	16:17:47.16	-53:23:23.0615	16.09	1.68	0.24	0.3603	EW
UCNS-V-923	5F	16:17:27.67	-53:31:43.2577	19.61	2.81	1.03	0.3122	EW
UCNS-V-935	5G	16:17:58.58	-53:28:30.9711	19.54	2.40	0.82	0.3566	EW
UCNS-V-941	5G	16:17:53.57	-53:27:48.4901	18.66	2.58	0.84	0.2410	EW
UCNS-V-944	5G	16:17:46.51	-53:27:35.2609	18.40	2.54	0.84	0.3445	EW
UCNS-V-1008	6B	16:18:26.62	-53:39:41.9407	16.17	1.50	0.60	0.3095	EW
UCNS-V-1060	6E	16:18:52.67	-53:33:40.0970	19.34	2.31	0.89	0.3187	EW
UCNS-V-1063	6E	16:18:42.36	-53:33:26.3744	16.28	1.47	0.65	0.3555	EW
UCNS-V-1066	6E	16:18:41.01	-53:32:31.6337	20.30	2.69	1.08	0.3732	EW
UCNS-V-1154	7B	16:19:27.62	-53:40:41.6333	17.22	1.82	0.87	0.4547	EW
UCNS-V-1199	7D	16:19:14.08	-53:37:10.5505	17.82	1.84	0.41	0.2723	EW
UCNS-V-1409	8F	16:20:16.80	-53:33:40.5853	16.49	1.31	0.50	0.4010	EW
UCNS-V-1417	8F	16:20:31.49	-53:33:04.9446	19.53	1.54	1.08	0.2378	EW

Table 6.5: Table of contact eclipsing binaries with maxima of different brightness.

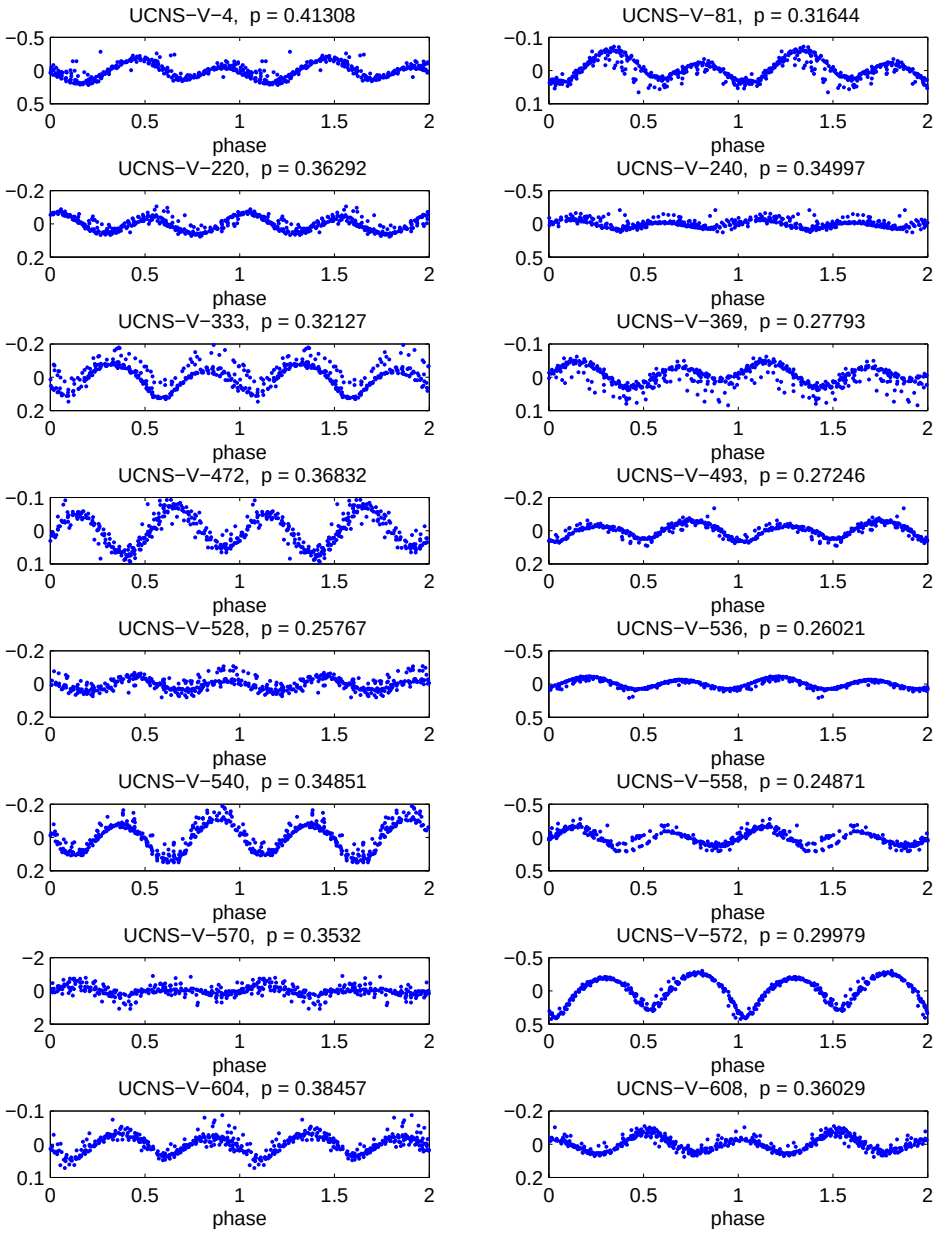


Figure 6.13: **Contact eclipsing binaries with two maxima (a)** Lightcurves of contact eclipsing binaries with maxima of different brightness.

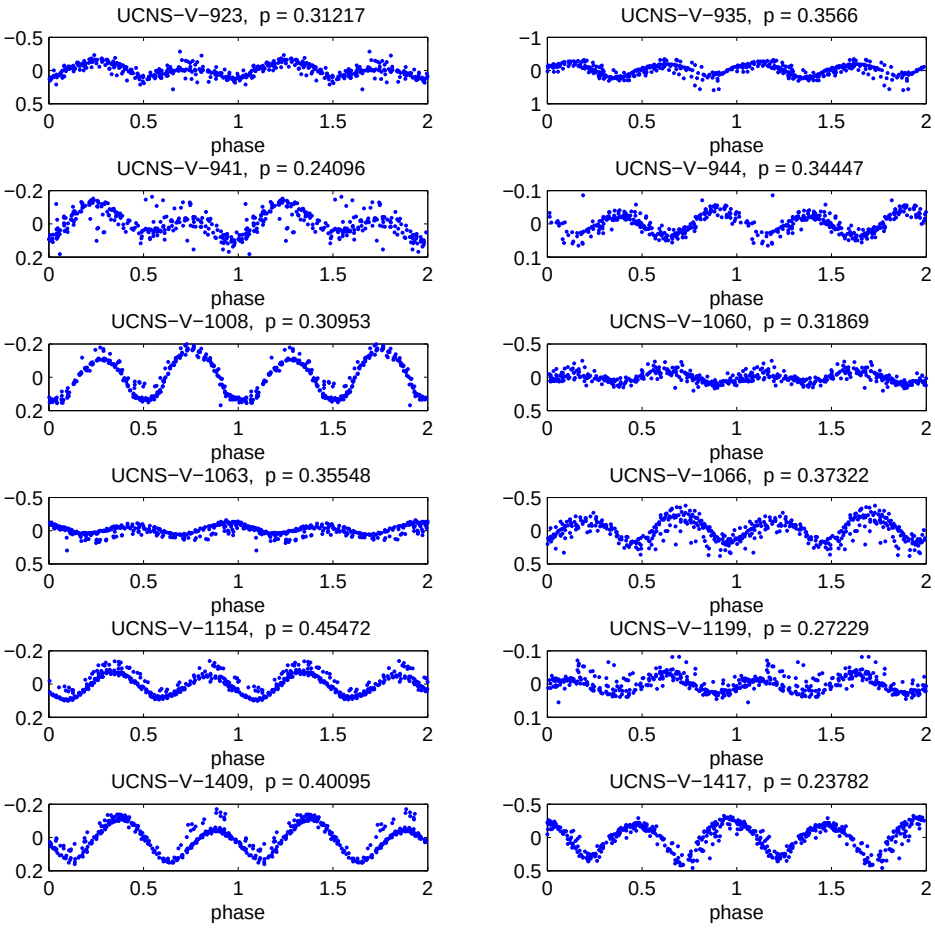


Figure 6.14: **Contact eclipsing binaries with two maxima (b)** Lightcurves of contact eclipsing binaries with maxima of different brightness.

6.4.1.1 Short period contact eclipsing binaries

Contact eclipsing binaries have a strong period cut-off at approximately 0.215 to 0.22 days with a population maximum at slightly longer periods - around 0.27 days. The period cutoff is well known but due to small numbers of close eclipsing binaries at short periods the reasons for the cutoff are unknown. Rucinski (2007) has made a study of contact binaries found in the All-Sky Automated Survey. The survey contains more than 3000 contact binaries but even with this large number the statistics at short periods are inconclusive. Any improvement in the numbers of contact binaries with periods less than 0.22 days is important. The shortest known period of a contact eclipsing binary is 0.2178 days (Rucinski, 2007).

From the variable stars found in this survey we find 3 variables identified as contact binaries with periods of less than 0.22 days, and another 4 with periods less than 0.23 days. Table 6.6 lists the three shortest period binaries and their lightcurves are shown in Figure 6.15. Unfortunately all three systems are faint and thus do not have very good lightcurves. Followup photometry on the systems is desired to confirm their nature and verify the periods. Two of the binaries have shorter periods that any published thus far.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-737	4G	16:17:50.67	-53:13:47.9457	19.19	2.65	1.16	0.2109	EW
UCNS-V-301	2F	16:19:27.64	-53:16:24.3907	22.14		1.68	0.2143	EW
UCNS-V-1085	6F	16:18:14.03	-53:32:50.9721	21.21	2.80	1.33	0.2199	EW

Table 6.6: **Short period contact eclipsing binaries** Table of contact eclipsing binaries with periods close to that of the known period cutoff.

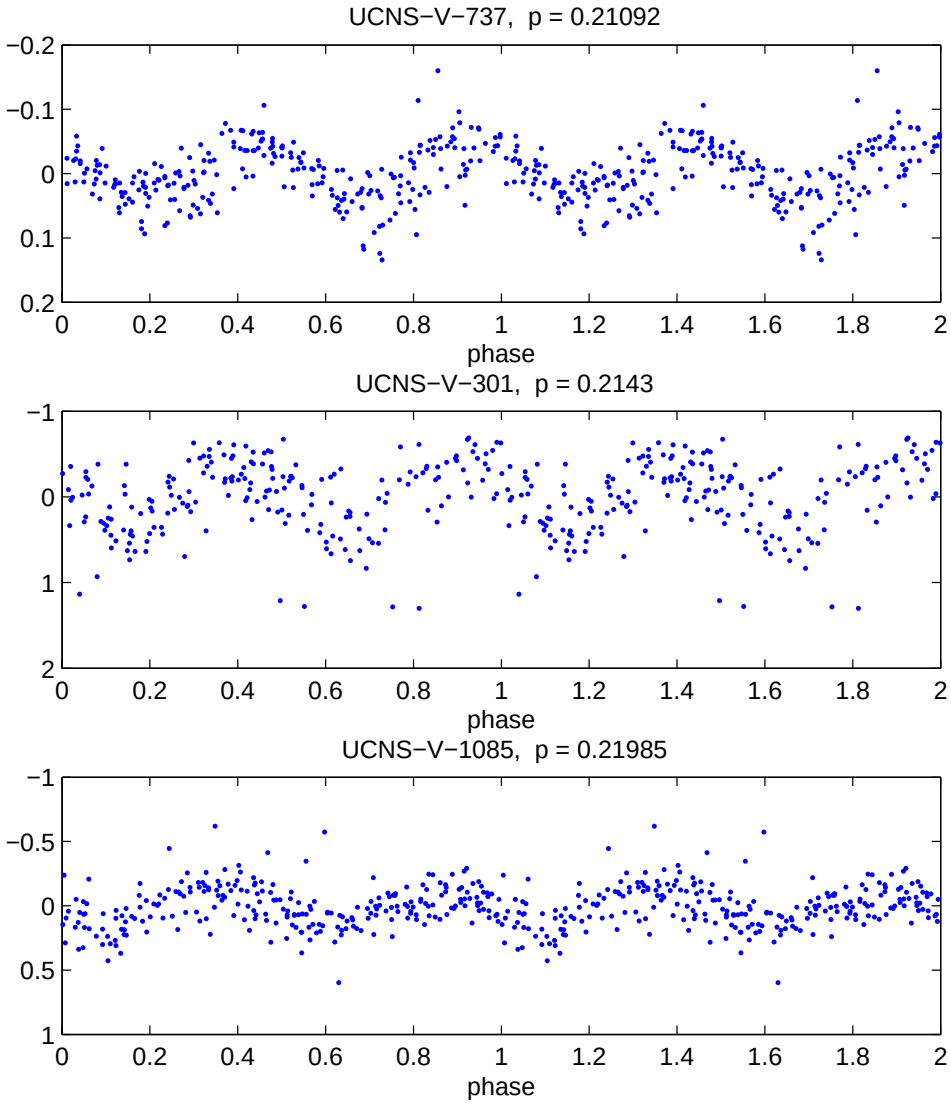


Figure 6.15: Lightcurves of the three detected contact eclipsing binaries with periods of less than 0.22 days.

6.4.1.2 Possible low mass-ratio contact eclipsing binaries

The literature on contact eclipsing binaries contains a number with low mass-ratios. The lightcurves of these systems always have a flat bottomed minimum and period within the range 0.3 to 0.4 d. Our catalogue of contains 18 contact binaries whose lightcurves correspond to these criteria (see Table 6.7 and Figure 6.16). There are also 14 contact binaries with flat bottomed lightcurves which fall outside the period range - 0.3 to 0.4d, (see Table 6.8 and Figure 6.17). The flat bottom of the minimum is caused by the full eclipse of the smaller component of the binary. Future observations of these variables could confirm the nature of their mass-ratios and add to the database.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-26	1A	16:20:43.91	-53:23:29.7361	19.17	2.27	1.96	0.3631	EW
UCNS-V-175	1H	16:20:08.57	-53:11:51.7763	17.98	2.34	0.81	0.3982	EW
UCNS-V-193	2A	16:19:38.23	-53:24:59.8532	16.47	1.25	0.97	0.3907	EW
UCNS-V-231	2C	16:19:42.23	-53:20:23.1327	18.42	2.30	0.48	0.3802	EW
UCNS-V-410	3B	16:18:22.56	-53:24:10.8596	19.03	1.69	1.00	0.3794	EW
UCNS-V-540	3G	16:18:50.54	-53:13:15.8250	18.10	2.13	0.52	0.3485	EW
UCNS-V-568	3H	16:18:15.50	-53:12:36.5601	17.37	2.03	0.66	0.3559	EW
UCNS-V-603	4A	16:18:04.46	-53:24:25.2664	15.73	1.92	0.47	0.3762	EW
UCNS-V-766	4H	16:17:31.19	-53:13:13.1713	17.07	1.93	0.51	0.3761	EW
UCNS-V-786	4H	16:17:23.08	-53:10:55.8408	18.15	2.34	0.71	0.3237	EW
UCNS-V-948	5G	16:18:04.49	-53:27:07.6592	17.07	2.09	0.05	0.3804	EW
UCNS-V-962	5H	16:17:27.09	-53:28:59.5381	19.12	2.30	0.62	0.3090	EW
UCNS-V-968	5H	16:17:38.22	-53:27:28.1100	16.90	1.64	0.59	0.3760	EW
UCNS-V-976	6A	16:19:00.14	-53:41:37.2113	17.33	2.45	0.04	0.3441	EW
UCNS-V-1117	6H	16:18:36.27	-53:28:26.7849	19.44	2.35	0.56	0.3212	EW
UCNS-V-1230	7E	16:19:56.70	-53:32:24.1384	15.89	1.06	0.38	0.3066	EW
UCNS-V-1350	8C	16:20:34.33	-53:37:18.2836	18.30	2.13	0.56	0.3506	EW
UCNS-V-1453	8G	16:20:45.64	-53:27:00.6708	18.61	2.40	0.17	0.3026	EW

Table 6.7: **Probable low mass-ratio contact eclipsing binaries** Table of contact eclipsing binaries which may have a low mass-ratio.

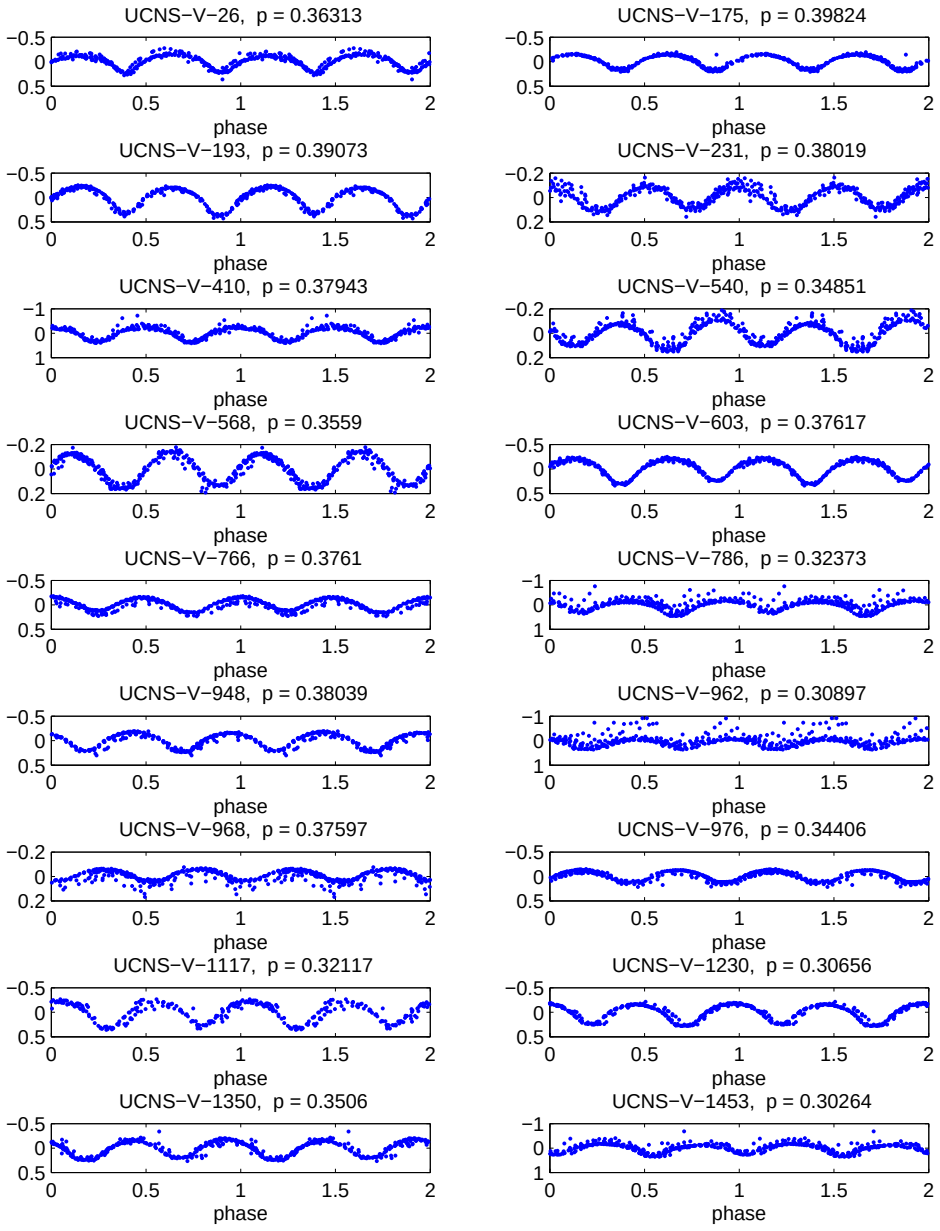


Figure 6.16: Lightcurves of the contact eclipsing binaries with shape and periods suggesting low mass-ratios.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-9	1A	16:20:50.50	-53:24:40.1987	19.05	1.33	2.28	0.5106	EW
UCNS-V-92	1C	16:20:48.80	-53:19:04.3178	19.24	2.23	0.94	0.4076	EW
UCNS-V-115	1E	16:20:53.66	-53:17:11.7310	19.14	2.55	0.67	0.5553	EW
UCNS-V-467	3E	16:18:59.97	-53:16:25.5253	19.25	2.83	0.52	0.4095	EW
UCNS-V-482	3E	16:18:56.66	-53:15:05.8831	17.15	1.95	0.53	0.4277	EW
UCNS-V-493	3F	16:18:36.50	-53:17:09.0202	16.23	1.87	1.75	0.2725	EW
UCNS-V-523	3F	16:18:30.89	-53:15:11.5286	16.49	2.38		0.9037	EW
UCNS-V-984	6A	16:18:47.74	-53:41:03.6335	17.74	2.84	0.61	0.6118	EW
UCNS-V-1091	6F	16:18:38.73	-53:31:56.7975	17.24	2.20	1.08	0.5765	EW
UCNS-V-1232	7E	16:19:39.57	-53:32:12.9626	18.13	1.65	0.90	0.4570	EW
UCNS-V-1272	7G	16:19:38.92	-53:27:04.6993	18.47	2.33	0.82	0.2853	EW
UCNS-V-1312	8A	16:20:38.43	-53:38:56.5849	17.00	1.24	4.28	0.5868	EW
UCNS-V-1354	8C	16:20:48.55	-53:36:20.9116	18.64	1.86	0.80	0.4498	EW
UCNS-V-1422	8F	16:20:09.11	-53:32:17.6392	18.48	1.92	0.75	0.4381	EW

Table 6.8: **Possible low mass-ratio contact eclipsing binaries** Table of contact eclipsing binaries which may have a low mass-ratio - the lightcurves have shape but not periods suggesting low mass-ratios.

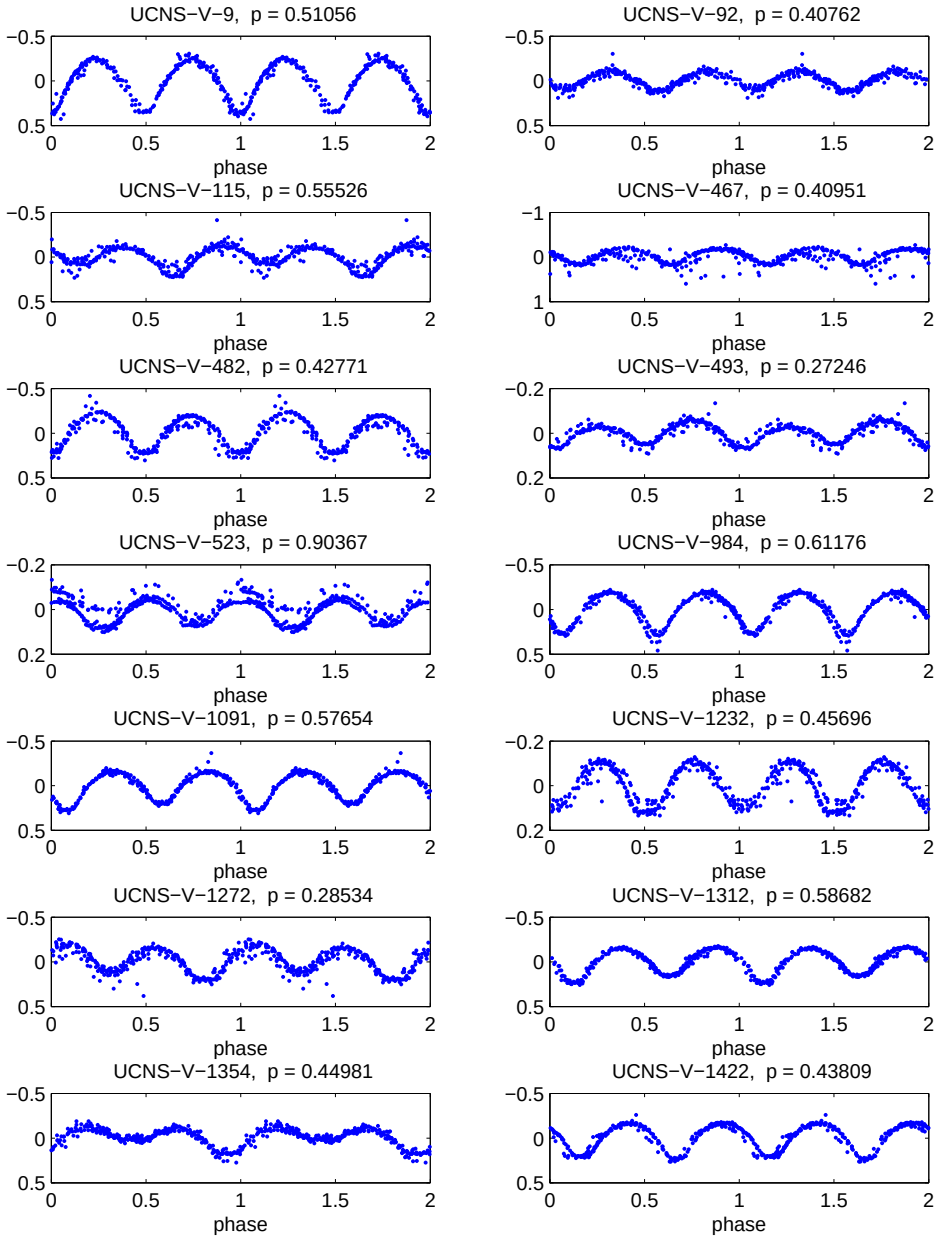


Figure 6.17: Lightcurves of the contact eclipsing binaries with shape (but not periods) suggesting low mass-ratios.

6.4.1.3 Low-mass component contact eclipsing binaries

As for eclipsing binaries with low mass-ratios it is possible to use the lightcurves found in the catalogue to identify possible systems with low-mass components. Observations of low-mass components in eclipsing binaries are important in order to obtain the stellar mass and radius which are then used in constraining stellar formation and evolution theory. As in Weldrake et al. (2007b), we select contact eclipsing binaries with periods of less than 0.25 days (23 candidates) and those Agol-type detached binaries with periods less than 1.6 days and non-varying out of eclipse lightcurves (9 candidates). The variables are listed in Tables 6.9 and 6.10 and their lightcurves shown in Figures 6.18 and 6.19. A number of these are brighter objects which would allow spectroscopic followup to confirm the nature of the components.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-737	4G	16:17:50.67	-53:13:47.9457	19.19	2.65	1.16	0.2109	EW
UCNS-V-301	2F	16:19:27.64	-53:16:24.3907	22.14		1.68	0.2143	EW
UCNS-V-1085	6F	16:18:14.03	-53:32:50.9721	21.21	2.80	1.33	0.2199	EW
UCNS-V-522	3F	16:18:28.65	-53:15:13.7680	21.54	3.46	2.35	0.2226	EW
UCNS-V-1250	7G	16:19:51.13	-53:29:59.1415	21.06	2.62	1.65	0.2234	EW
UCNS-V-840	5C	16:17:57.44	-53:36:59.9529	21.49	3.02	1.68	0.2270	EW
UCNS-V-1386	8D	16:20:07.71	-53:35:31.1146	20.01	1.91	2.30	0.2289	EW
UCNS-V-179	1H	16:20:19.37	-53:11:35.1909	20.51	3.03	1.13	0.2321	EW
UCNS-V-1417	8F	16:20:31.49	-53:33:04.9446	19.53	1.54	1.08	0.2378	EW
UCNS-V-776	4H	16:17:34.08	-53:12:02.1376	19.56	2.53	0.97	0.2386	EW
UCNS-V-1056	6D	16:18:15.23	-53:35:25.1012	21.87	2.96	1.21	0.2389	EW
UCNS-V-941	5G	16:17:53.57	-53:27:48.4901	18.66	2.58	0.84	0.2410	EW
UCNS-V-434	3C	16:18:54.81	-53:20:48.4850	20.18	2.62	0.66	0.2413	EW
UCNS-V-740	4G	16:17:49.64	-53:13:21.5601	17.91	2.34	0.96	0.2429	EW
UCNS-V-752	4G	16:17:56.26	-53:11:51.7213	17.88	1.90	0.76	0.2459	EW
UCNS-V-754	4G	16:17:56.45	-53:11:50.6632	19.94	2.67	1.15	0.2459	EW
UCNS-V-753	4G	16:17:56.34	-53:11:51.1016	18.01	2.20	0.40	0.2460	EW
UCNS-V-558	3G	16:18:54.13	-53:11:08.7186	18.95	2.33	0.85	0.2487	EW
UCNS-V-1434	8F	16:20:20.06	-53:31:01.2833	20.47	2.48	0.89	0.2490	EW
UCNS-V-210	2A	16:19:57.61	-53:22:41.2599	19.85	2.22	1.48	0.2493	EW
UCNS-V-516	3F	16:18:31.63	-53:15:35.7031	19.64	3.42	2.40	0.2493	EW
UCNS-V-1446	8G	16:20:51.24	-53:27:48.3616	17.33	2.00	0.32	0.2494	EW
UCNS-V-1447	8G	16:20:51.11	-53:27:48.4544	17.43	2.31	0.40	0.2494	EW

Table 6.9: **Low mass component contact eclipsing binaries** Table of eclipsing binaries which may have a low mass component.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-1106	6G	16:18:55.35	-53:28:34.7387	18.83	2.19	0.73	0.4983	EA
UCNS-V-992	6A	16:18:45.09	-53:39:45.9131	19.46	2.46	1.48	0.4095	EA
UCNS-V-790	4H	16:17:28.84	-53:10:42.0607	17.83	2.09	0.37	0.4711	EA
UCNS-V-244	2C	16:19:37.54	-53:18:34.5958	18.35	2.84	1.14	0.5610	EA
UCNS-V-755	4G	16:17:58.34	-53:11:39.78	20.63	3.00	1.28	0.6861	EA
UCNS-V-611	4A	16:18:02.91	-53:23:07.2723	18.33	2.37	0.50	0.7061	EA
UCNS-V-979	6A	16:18:46.94	-53:41:25.8984	18.57	2.91	0.43	0.8337	EA
UCNS-V-894	5E	16:18:05.07	-53:31:34.5940	19.24	1.60	1.65	0.7043	EA
UCNS-V-52	1B	16:20:11.65	-53:24:02.0859	16.28	1.86	0.19	1.1242	EA

Table 6.10: **Low mass component detached eclipsing binaries** Table of eclipsing binaries which may have a low mass component.

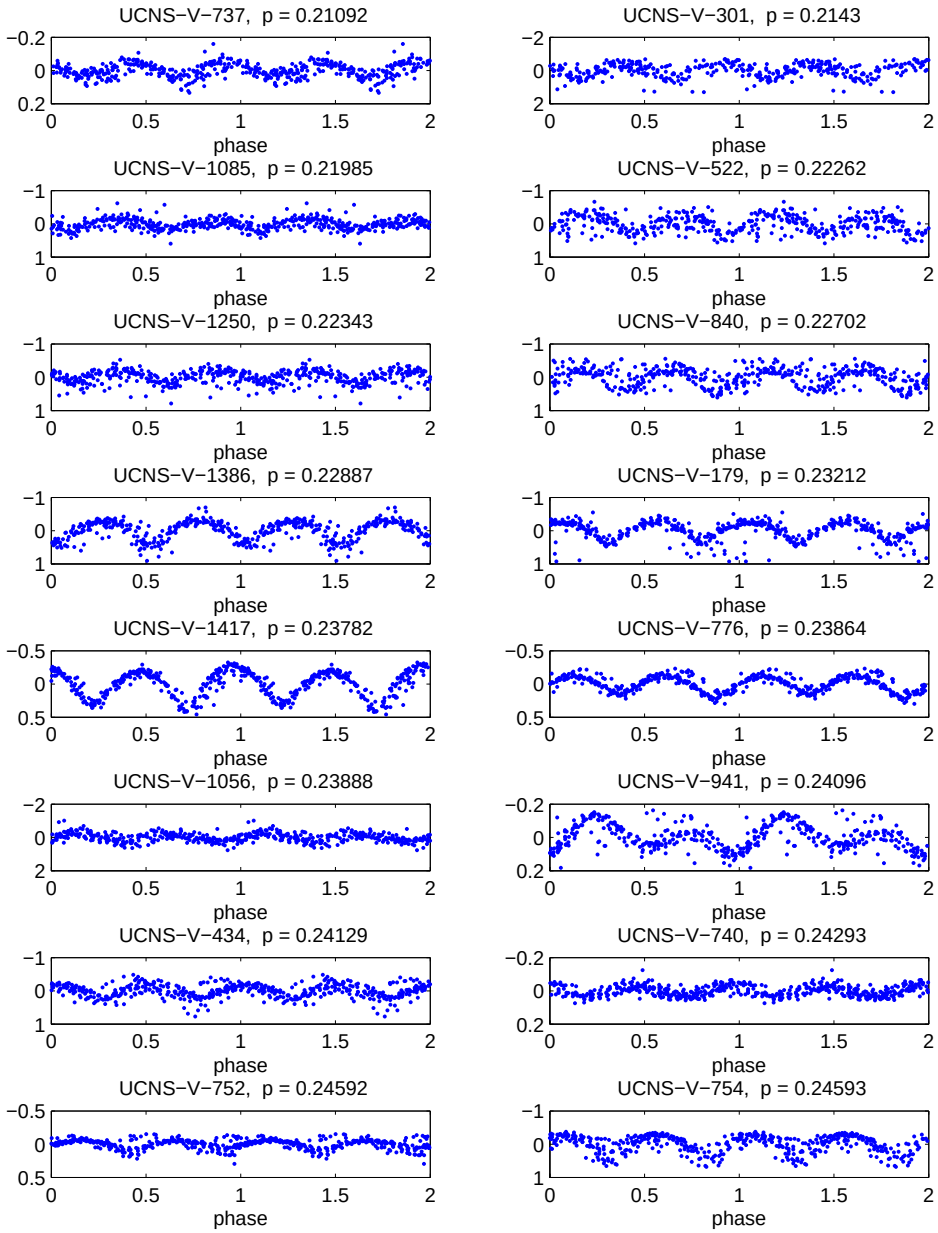


Figure 6.18: **Lightcurves of eclipsing binaries with possible low-mass components (a)** Lightcurves of eclipsing binaries with shape and periods suggesting a low-mass component.

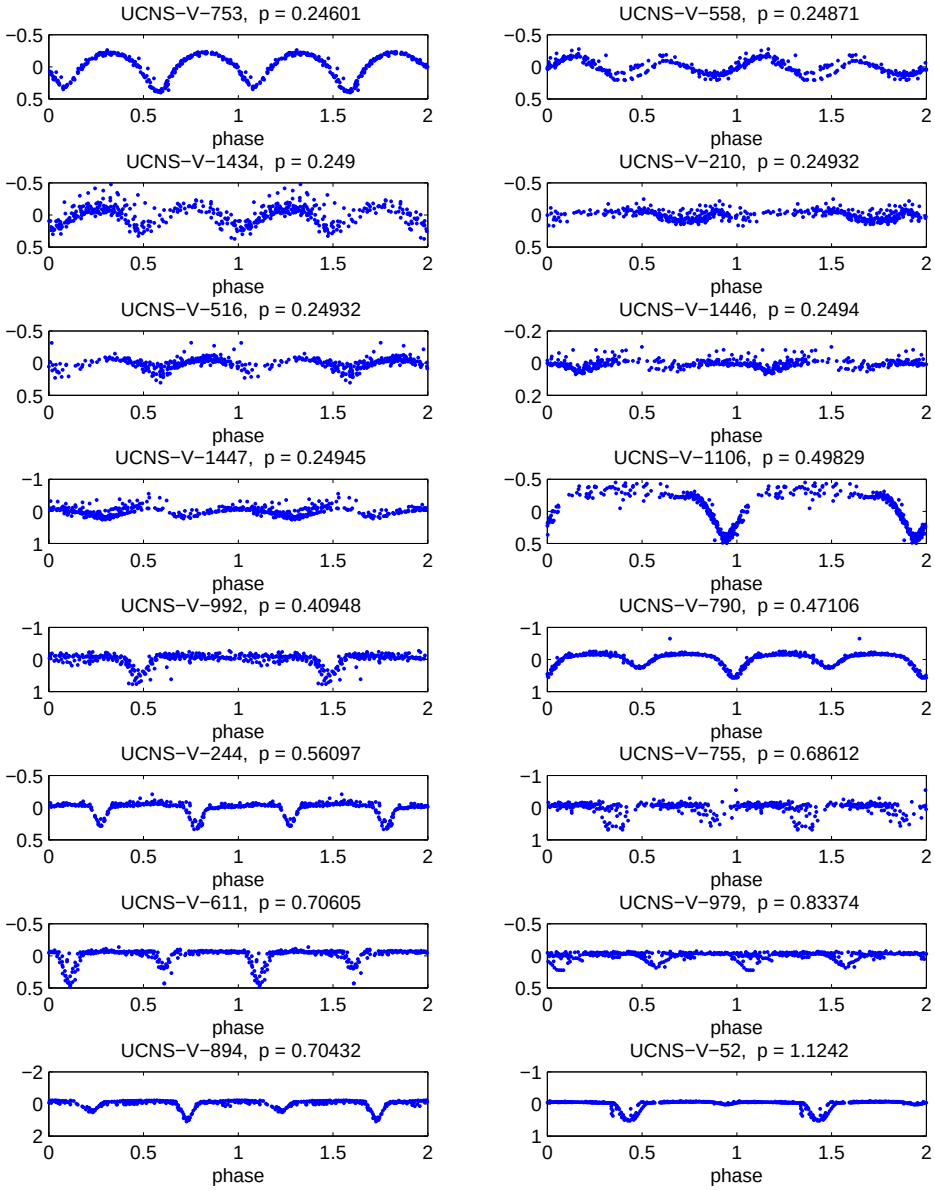


Figure 6.19: **Lightcurves of eclipsing binaries with possible low-mass components (b)** Lightcurves of the eclipsing binaries with shape and periods suggesting a low-mass component.

6.4.2 Variable stars off the main sequence

Examination of the variable stars which appear off the main sequence in the colour-magnitude diagrams has not revealed any particular trends. There are a large range of periods present and the variables do not all appear as one type. There are a number of stars with a period close to one day or an alias of such and this fraction seems higher than that in the total catalogue. The off main-sequence stars are tabulated in Table 6.11, with the redder stars listed first then the bluer. The lightcurves of these variables can be seen in Figures 6.20 to 6.24. Cross referencing the stars found off the main sequence with those discussed elsewhere in this chapter have found two previously mentioned. The first of these is a contact binary (UCNS-V-1230) which appears redder than the main sequence. This variable is also discussed in §6.4.1.2 and may have a low mass ratio. Eclipsing binaries (detached and semi-detached) which are bluer than the main sequence may contain a white-dwarf component. A number of the bluer stars are eclipsing binaries and one of these stars (UCNS-V-894) fits the criteria for a low mass companion (see §6.4.1.3). Figure 6.24 contains all those variable stars found to be bluer than the main sequence.

Table 6.11: **Variable stars off the main sequence** Table of variable stars which do not fall on the main sequence seen in the colour-magnitude diagrams.

ID	Subset	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-70	1C	16:20:42.84	-53:21:34.5504	17.13	2.76	0.97	19.0426	EB
UCNS-V-96	1C	16:20:51.92	-53:18:49.0723	17.95	2.99	1.01	0.9539	PUL:
UCNS-V-104	1D	16:20:14.62	-53:21:05.4916	17.58	2.11	0.75	8.0431	PUL:
UCNS-V-105	1D	16:20:12.64	-53:20:57.3738	16.27	3.34		1.9334	EB/PUL
UCNS-V-112	1D	16:20:19.04	-53:19:30.1081	20.02	3.23	1.63	0.6662	EW
UCNS-V-139	1G	16:20:43.97	-53:13:42.5724	16.18	4.40		0.9616	PUL:
UCNS-V-162	1G	16:20:45.36	-53:10:40.3784	16.25	4.19	0.90	1.0604	PUL:
UCNS-V-173	1H	16:20:24.63	-53:11:52.5340	20.93	4.13	1.14	0.3194	EW
UCNS-V-199	2A	16:19:47.02	-53:24:33.1889	18.12	2.74	1.53	1.9181	EB:
UCNS-V-202	2A	16:19:56.18	-53:23:54.6145	19.84	4.40	-2.18	2.2439	EB:
UCNS-V-411	3B	16:18:13.30	-53:24:09.7108	16.08	3.87		10.1541	PUL
UCNS-V-419	3B	16:18:36.16	-53:23:36.5147	16.00	3.90		0.9631	PUL:
UCNS-V-441	3C	16:18:48.65	-53:19:39.9446	16.09	2.66	0.90	13.0928	EB/PUL
UCNS-V-447	3D	16:18:31.89	-53:21:24.9115	17.58	2.69	1.67	2.1272	EB:
UCNS-V-525	3F	16:18:13.90	-53:15:08.4782	16.37	3.13	2.25	0.3211	EW
UCNS-V-581	3H	16:18:16.88	-53:11:26.9764	17.10	2.70	0.90	2.2406	EB:
UCNS-V-582	3H	16:18:24.83	-53:11:25.5501	15.99	2.84	1.08	17.1408	PUL:
UCNS-V-601	4A	16:17:55.06	-53:24:28.2647	16.06	3.29	0.81	0.9683	PUL
UCNS-V-605	4A	16:18:00.95	-53:24:03.3517	15.88	4.29		1.0447	PUL:
UCNS-V-750	4G	16:17:53.29	-53:12:02.1058	18.30	3.11	1.31	0.3793	EW
UCNS-V-771	4H	16:17:20.93	-53:12:58.9279	16.18	3.72		0.9648	PUL
UCNS-V-828	5B	16:17:21.64	-53:39:19.0873	17.80	3.38	0.87	1.0408	PUL:
UCNS-V-832	5C	16:17:53.30	-53:37:56.4531	18.73	3.71	1.13	4.3532	EB:

Continued on next page

Table 6.11 – continued from previous page

ID	Subset	Right Ascension	Declination	Calibrated Magnitudes			Period	Type
	(J2000.0)	(J2000.0)	R	B-R	R-I	(days)		
UCNS-V-837	5C	16:18:02.56	-53:37:11.4579	18.68	3.19	1.17	0.2926	EW
UCNS-V-848	5C	16:17:50.66	-53:35:50.1725	20.40	4.25	1.36	3.4134	EB:
UCNS-V-854	5D	16:17:39.27	-53:37:57.1063	19.61	4.41	0.86	0.6904	PUL:
UCNS-V-870	5D	16:17:16.83	-53:35:05.0840	15.46	3.25	-0.11	1.6404	CV/PUL
UCNS-V-930	5G	16:17:43.06	-53:29:08.6007	15.97	4.31		0.9319	PUL:
UCNS-V-954	5H	16:17:18.38	-53:29:28.6719	17.37	2.93	1.05	9.4259	PUL:
UCNS-V-1019	6C	16:18:46.31	-53:37:53.8736	18.02	2.79	0.77	2.1220	EB/PUL
UCNS-V-1027	6C	16:18:41.87	-53:36:27.6245	17.42	2.46	0.80	5.3830	EB:
UCNS-V-1035	6C	16:18:47.72	-53:35:47.0940	17.70	2.67	0.95	19.7144	EB:
UCNS-V-1041	6D	16:18:34.40	-53:37:25.3713	17.91	3.16	1.14	6.3022	EB
UCNS-V-1050	6D	16:18:18.11	-53:36:29.9025	16.50	2.49	0.88	0.4408	EW
UCNS-V-1083	6F	16:18:31.07	-53:33:08.9237	18.04	2.79	1.49	5.1341	PUL:
UCNS-V-1089	6F	16:18:33.25	-53:32:13.2173	17.98	2.99	1.30	1.8838	EB:
UCNS-V-1125	7A	16:19:51.01	-53:41:34.1022	18.04	3.10	0.74	0.3636	EW
UCNS-V-1126	7A	16:19:44.41	-53:41:26.9105	17.18	2.82	0.59	0.9238	PUL
UCNS-V-1174	7C	16:19:40.58	-53:37:16.9534	17.55	2.50	2.06	0.9244	EW:
UCNS-V-1190	7C	16:19:37.75	-53:35:11.8051	17.85	3.08	1.31	1.0495	PUL
UCNS-V-1234	7E	16:19:45.10	-53:31:47.7425	18.39	2.29	1.14	0.4646	EW
UCNS-V-1230	7E	16:19:56.70	-53:32:24.1384	15.89	1.06	0.38	0.3066	EW
UCNS-V-1233	7E	16:19:46.19	-53:31:51.0483	16.36	1.61	0.85	0.4105	EW
UCNS-V-1235	7E	16:19:43.60	-53:31:45.7793	16.11	0.73	0.24	0.3249	EW
UCNS-V-1241	7F	16:19:29.61	-53:33:46.4558	16.56	2.91	1.04	0.9557	PUL
UCNS-V-1244	7F	16:19:19.41	-53:33:13.8689	17.80	2.84	1.02	1.0914	PUL
UCNS-V-1294	7H	16:19:25.51	-53:27:18.8946	16.33	2.32	0.64	5.6922	DCEP/PUL
UCNS-V-1300	8A	16:20:44.19	-53:41:13.1280	16.54	2.76	4.33	0.9572	PUL
UCNS-V-1348	8C	16:20:56.72	-53:37:33.4913	16.37	2.15	0.68	7.8091	PUL:

Continued on next page

Table 6.11 – continued from previous page

ID	Subset	Right Ascension	Declination	Calibrated Magnitudes			Period	Type
	(J2000.0)	(J2000.0)	R	B-R	R-I	(days)		
UCNS-V-1349	8C	16:20:46.79	-53:37:23.9814	17.48	2.63	0.88	2.0982	EB:
UCNS-V-1361	8C	16:20:43.01	-53:35:12.0040	16.56	2.21	0.89	2.1402	EB
UCNS-V-1362	8C	16:20:56.13	-53:35:10.4319	17.50	2.63	1.00	9.2446	EB:
UCNS-V-1380	8D	16:20:16.12	-53:36:07.2759	15.95	2.52	2.00	7.8590	PUL
UCNS-V-149	1G	16:20:56.01	-53:11:41.6310	21.46	1.11	1.18	1.9232	EB:
UCNS-V-153	1G	16:20:48.70	-53:11:24.4095	22.08		1.18	2.0894	EB:
UCNS-V-161	1G	16:20:38.50	-53:10:42.5178	21.33	2.35	0.99	0.4007	EW:
UCNS-V-364	3A	16:19:01.65	-53:25:40.0235	21.38	2.24	1.29	0.2955	EW
UCNS-V-382	3A	16:18:54.21	-53:23:45.0522	21.91	1.84	0.73	0.3189	EW
UCNS-V-470	3E	16:18:41.70	-53:16:22.4156	21.87	1.68	0.27	0.9197	PUL:
UCNS-V-542	3G	16:18:52.29	-53:13:13.8844	20.12	1.82	0.69	0.3413	EW
UCNS-V-576	3H	16:18:18.61	-53:12:13.2789	21.19	2.05	0.91	0.3609	EW
UCNS-V-587	3H	16:18:21.26	-53:10:49.3220	21.55	2.32	1.32	0.3487	EW
UCNS-V-620	4B	16:17:36.00	-53:25:04.7264	18.35	1.81	0.63	0.3971	EW
UCNS-V-801	5A	16:18:01.48	-53:40:19.5726	20.59	1.83	0.34	0.6746	CV/PUL
UCNS-V-866	5D	16:17:17.11	-53:35:29.5915	18.43	2.18	0.34	3.2450	EA:
UCNS-V-880	5E	16:17:48.20	-53:32:53.2065	19.03	1.41	1.24	0.3325	EW/EA
UCNS-V-894	5E	16:18:05.07	-53:31:34.5940	19.24	1.60	1.65	0.7043	EA
UCNS-V-1102	6G	16:18:43.98	-53:29:36.7935	19.92	0.82	1.07	0.3207	EW
UCNS-V-1319	8B	16:20:20.87	-53:41:30.5543	21.63	0.70	1.79	2.2396	EB:

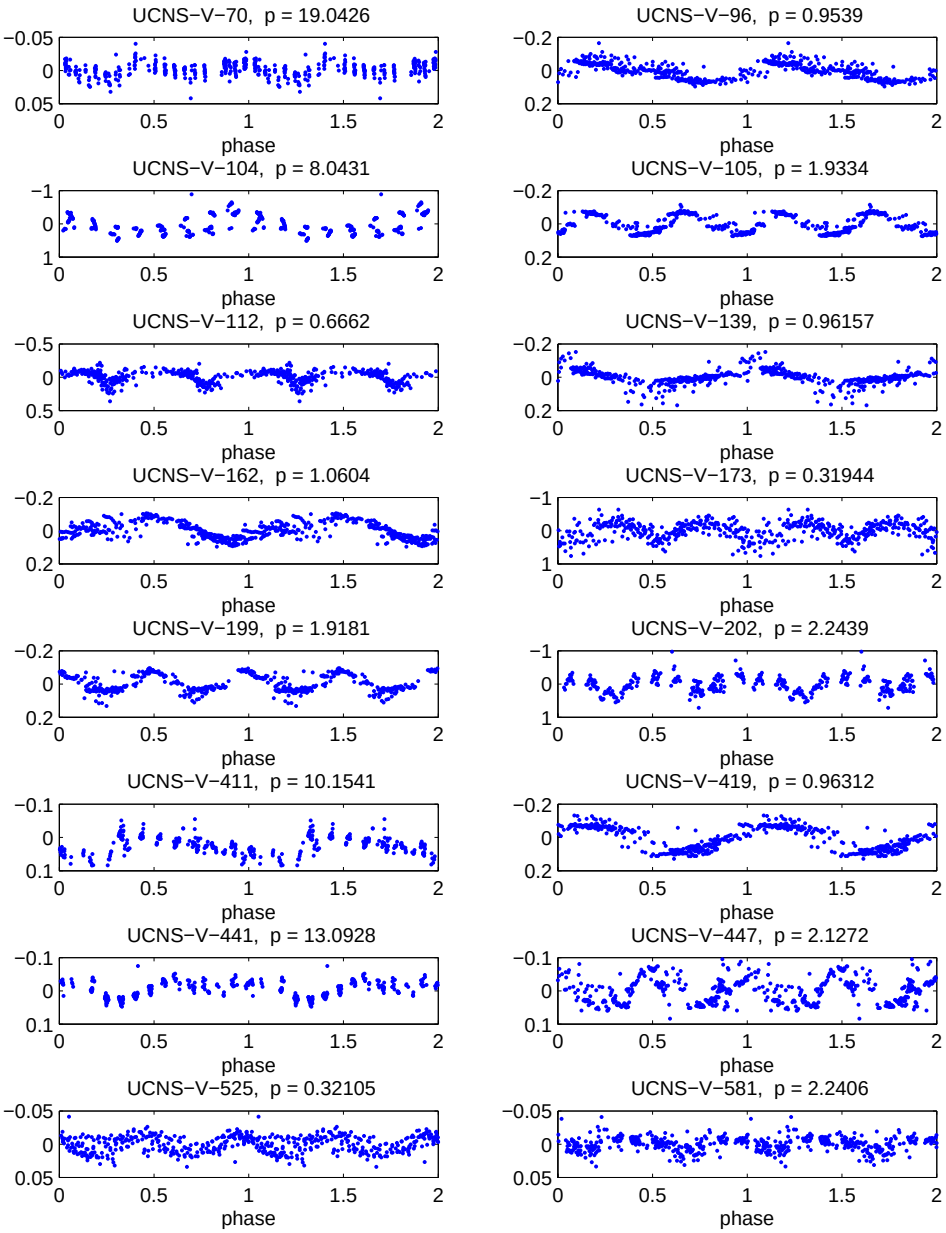


Figure 6.20: **Variable stars redder than the main sequence (a)** Lightcurves of variable stars which appear redder than the main sequence seen in the colour-magnitude diagrams.

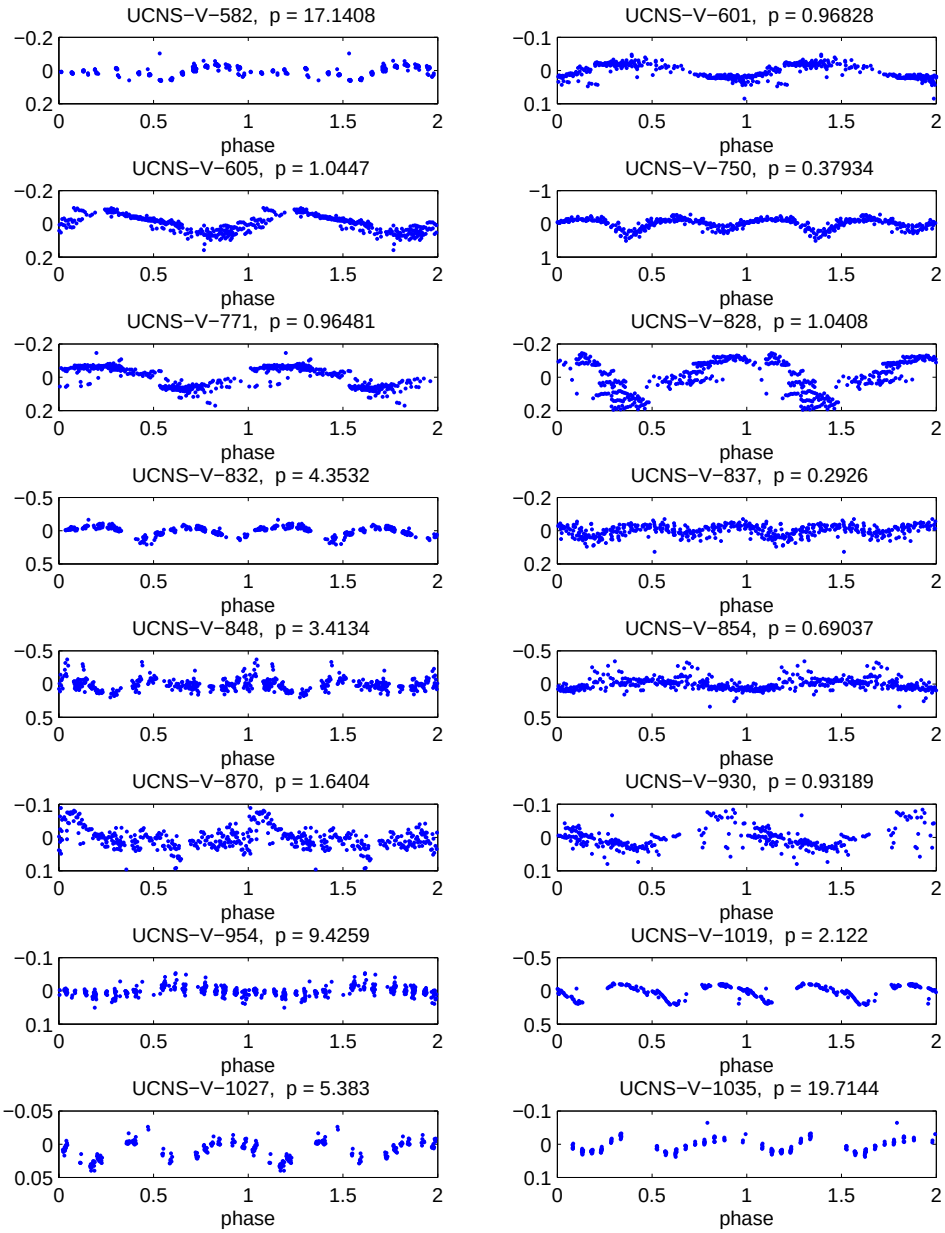


Figure 6.21: **Variable stars redder than the main sequence (b)** Lightcurves of variable stars which appear redder than the main sequence seen in the colour-magnitude diagrams.

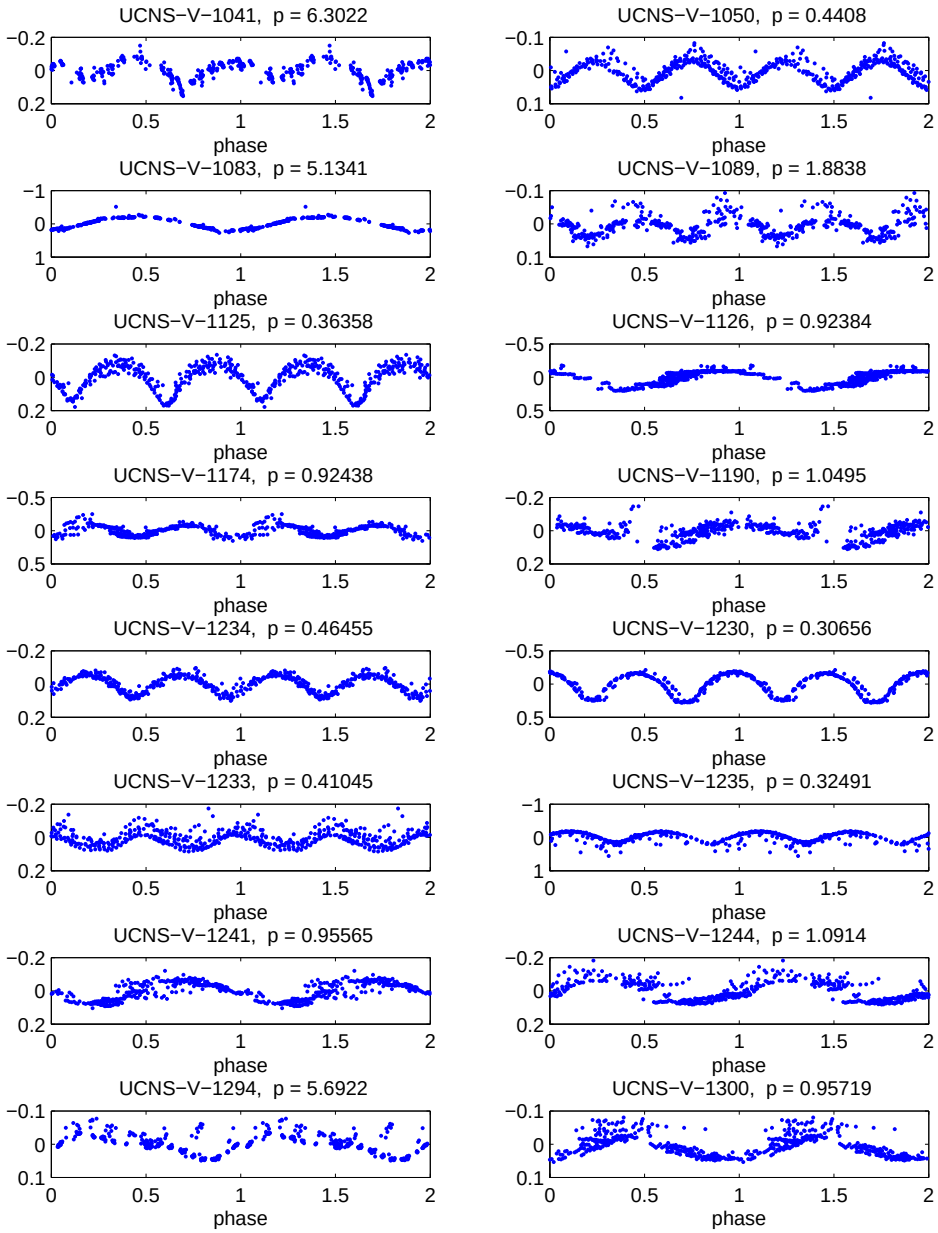


Figure 6.22: **Variable stars redder than the main sequence (c)** Lightcurves of variable stars which appear redder than the main sequence seen in the colour-magnitude diagrams.

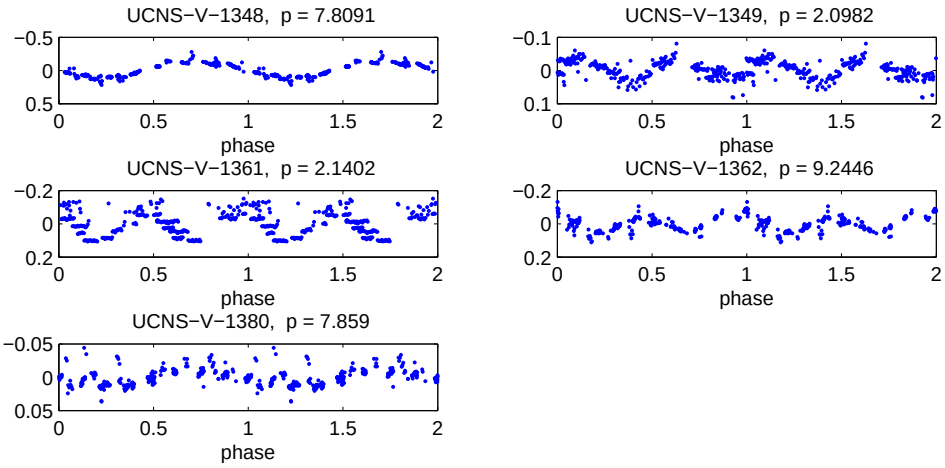


Figure 6.23: **Variable stars redder than the main sequence (d)** Lightcurves of variable stars which appear redder than the main sequence seen in the colour-magnitude diagrams.

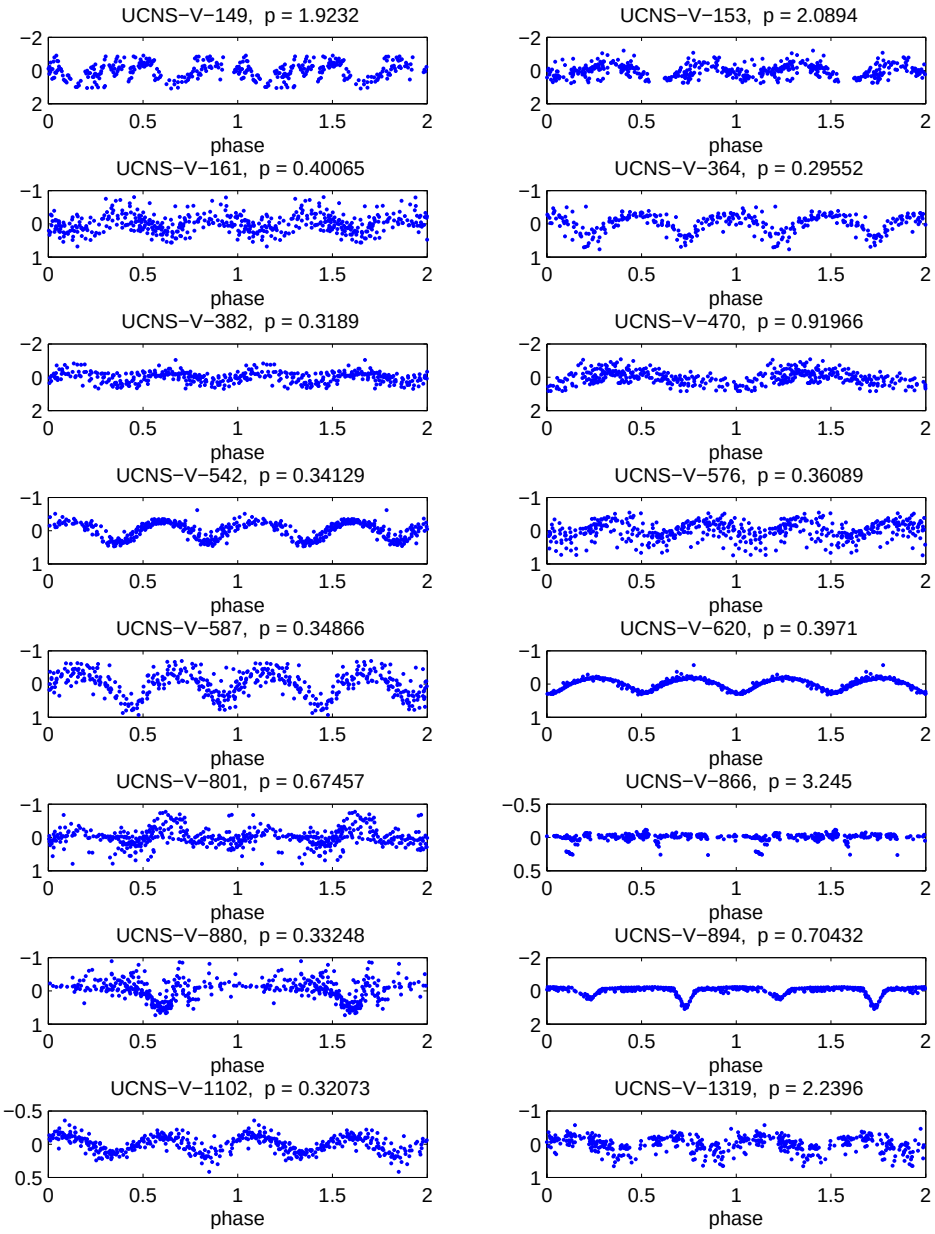


Figure 6.24: **Variable stars bluer the main sequence** Lightcurves of variable stars which appear bluer than the main sequence seen in the colour-magnitude diagrams.

6.4.3 Binary fraction

The total fraction of binary stars in the field can be calculated from the number of detected contact binaries. As the variable star detection efficiency is not 100% we must take this into account when we determine the binary fraction.

As discussed in §6.1.1, the variable detection efficiency drops at fainter magnitudes. From an average amplitude (150 mmag) we determine a slope of the detection efficiency and then determine an effective number of stars, Figure 6.25. The effective number of stars is thus calculated to be 62.60% of the total number of stars. The binary fraction is then calculated with a star count of 335592.

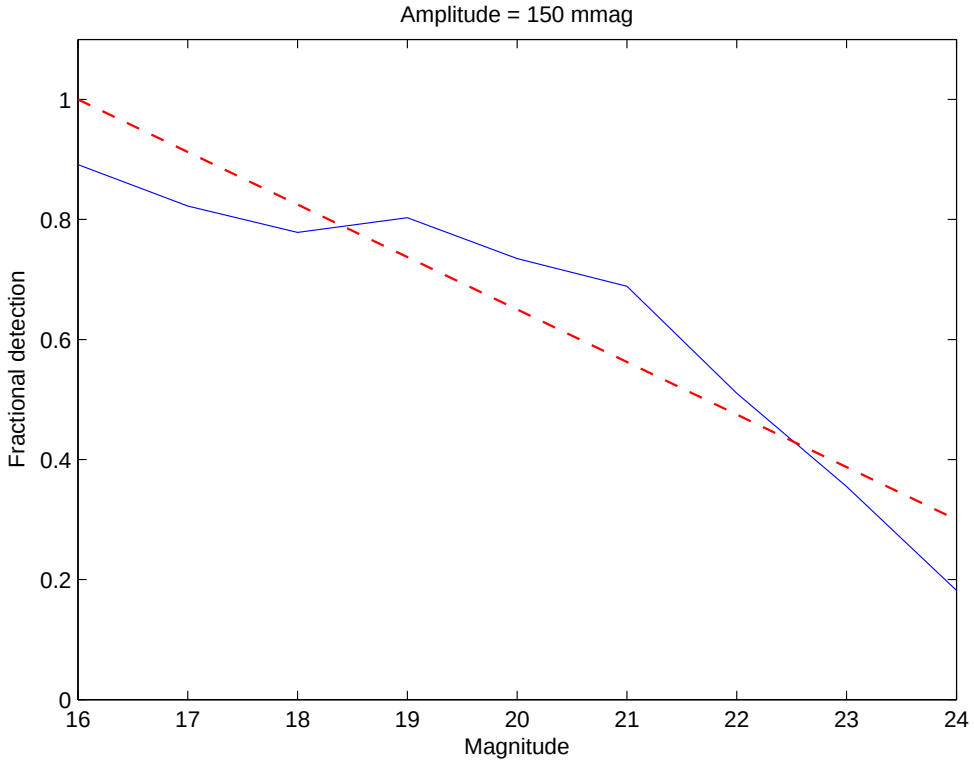


Figure 6.25: **Effective number of stars from the variable detection efficiency** Shown here are correctly detected variables as a fraction of inserted sinusoidal variables as a function of magnitude for 150 mmag amplitude simulated variables. The dashed line represents an average slope which is used to calculate the effective number of stars searched by the variable algorithm.

From the fraction of detected contact eclipsing binaries we can make an estimate of the total fraction of binary stars in the field. There are 588 detected contact eclipsing binaries in this field. The effective number of field stars is 335592 giving an observed fraction of contact binaries as 0.175%. Taking into account non-detections due to unfavourable orientations causes this number to double to 0.35%, which tallies well with the rate of contact binaries found amongst main-sequence stars in the disk of the Galaxy by Rucinski (1997).

The time taken for an established binary to merge and become a single star is on the order of $10^{9\pm1}$ yr (Vilhu, 1982; Rahunen, 1981). If we take 10^9 as the lifetime of a contact binary requires one contact binary formed every 8.5 x 10^5 yr. If we instead take the contact lifetime as 10^{10} rather than 10^9 the derived frequencies will drop by a factor of 10. We adopt here 10^9 yr as the lifetime of a contact binary and realise that because of the uncertainty in this figure the binary fractions we determine have associated large errors.

Having no turn-off in the colour-magnitude diagram of the field we are unable to determine an age for the main portion of the field. Instead we take three options and follow these to determine the binary fraction.

Age 1 Gyr:

Given an age of 1 Gyr for the field, 1176 contact binaries would have formed over the age of the field. At this age we find from Bradstreet and Guinan (1994) that only binaries with initial periods around 2.5 days will have come to contact. We find a primordial frequency of 0.35% for those binaries with initial periods around 2.5 days. Making the assumption that the period distribution is flat with log P (and using the period range 2.5 to 2.58 days) we determine a total primordial binary fraction of 100% $\pm$ 4% (where the error quoted here is simply the Poisson noise and in reality the error in the binary fraction is far larger).

Age 5 Gyr:

With an approximately solar age field (which is most likely from the location) we

find that 5882 contact binaries would have formed over the age of the field. From Bradstreet and Guinan (1994) we find that at this age those binaries that will have evolved to contact will have primordial periods between 2.5 and 3.5 days. If all such binaries have evolved to contact we find a primordial binary frequency of 1.75% for those with initial periods between 2.5 and 3.5 days. Again assuming the period distribution is flat with log P we determine a total primordial binary fraction of $46\% \pm 4\%$.

Age 10 Gyr:

At double the solar age (and close to the age of the globular cluster 47 Tucanae) we find that 11764 contact binaries would have formed over the age of the field. At this age Bradstreet and Guinan (1994) indicates that binaries forming to contact will have had initial periods between 2.5 and 4.0 days. Assuming all binaries of periods between 2.5 and 4.0 days have evolved to contact the numbers imply a primordial frequency of 3.5% with periods between 2.5 and 4.0 days. Assuming the period distribution is flat with log P we determine a total primordial binary fraction of $66\% \pm 4\%$.

To compare these values to that for stars in the solar neighbourhood, we take the 65% frequency and "exponential" period distribution used by Duquennoy and Mayor (1991). Considering binary stars with periods between 2.5 days and 50 yr the frequency amongst solar-type stars is 26%. Albrow et al. (2001), using the same methods as here, finds the binary fraction of 47 Tucanae (within 90" of the centre) as 14%, approximately half that of the solar field. Here, for a solar age field, we find a binary fraction from contact binaries of 46%, i.e. around 1.75 times larger than that found for nearby solar-type stars.

The binary fraction calculated here is poorly constrained due to the large uncertainty in the binary merger timescale and the uncertainty in the age of the field. The binary fractions are quoted here with statistical errors. The errors calculated from our uncertainties in values used in the calculations are far larger.

Ideally we would also make an estimate of the binary fraction from the detached binaries found in the field. However the method we have for this (Albrow et al., 2001) requires the mass and radii of the detached binary components. Without a turn-off on the colour-magnitude diagrams we are unable to fit isochrones to determine these.

6.4.4 Pre-main sequence stars

Detached eclipsing binaries are important for constraining stellar evolutionary theories. From detached eclipsing binaries one can determine elementary stellar information such as the mass and radius. Calibration of pre-main sequence (PMS) stars is scarce below 1 solar mass. There are only 6 PMS detached eclipsing binaries known in this mass range (Irwin et al., 2007a) and fitting of these by current models fails when both components are fitted simultaneously. This implies that current evolutionary models need revision. Detections of more pre-main sequence detached binaries will lead to more knowledge of the area and stronger constraints on the theory. Following the method used in Christiansen et al. (2008b), we use the positions of known pre-main sequence binaries on a colour-colour diagram, to determine which detached binaries in a sample may be pre-main sequence. Those can then be followed up by spectroscopy (Irwin et al., 2007a), to search for pre-main sequence signatures such as Lithium lines to confirm the nature of the binary stars.

In order for the comparison to be made, J,H and K magnitudes need to be obtained from the 2MASS survey (Skrutskie et al., 2006) for the detached binaries in this survey so that they can be compared to the known pre-main sequence binaries.

Due to the faintness of many of the stars in the field only 3 of the EA type stars (UCNS-V-827, UCNS-V-887 and UCNS-V-897) were given magnitudes by the 2MASS survey though the variability was not detected by this survey. A colour-colour diagram of the 6 known PMS eclipsing binaries and our three EA stars with 2MASS magnitudes is plotted in Figure 6.26. The dashed lines represent limits on candidature of pre-main sequence binaries (J - H = 0.6 mag, H - K = 0.078 mag). Also plotted are the intrinsic stellar loci of giant stars (••) and dwarf stars (•−) from Bessell and Brett (1988). Two of our EA stars (UCNS-V-827 and UCNS-V-887) are possible candidates for pre-main sequence stars and should therefore be followed up further, Table 6.12. Given that 2 of 3 detached binaries were found to be possible pre-main sequence stars, further deep infrared observations would be

desirable to determine if other detached binaries found in our survey also fall into the pre-main sequence area of the colour-colour diagram.

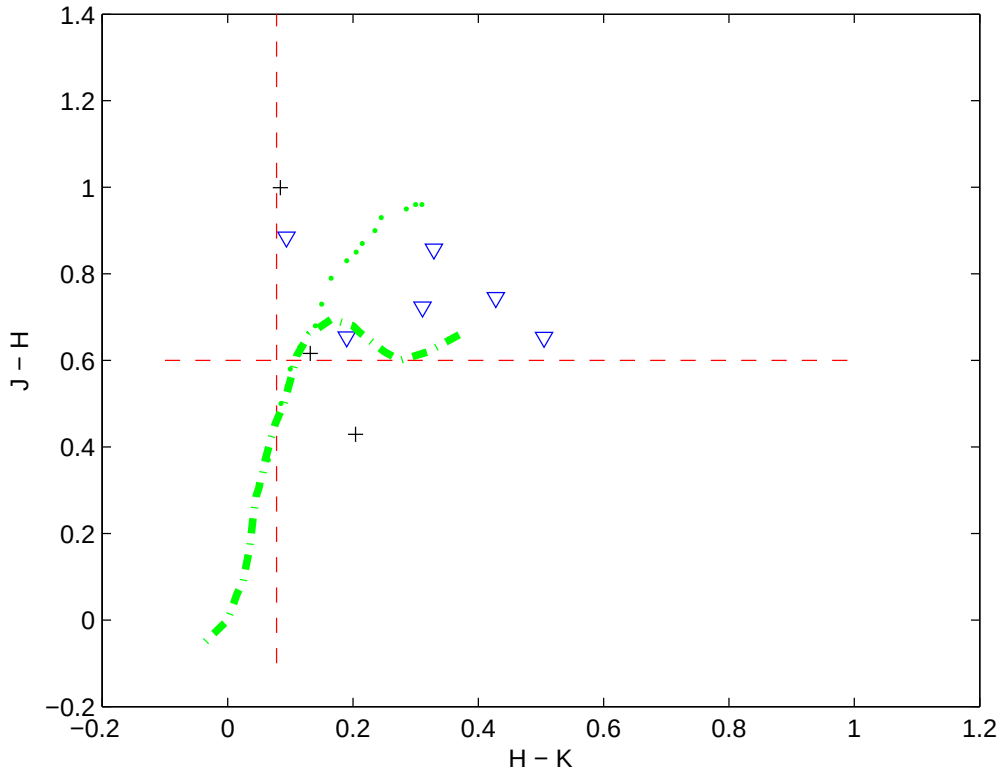


Figure 6.26: Colour-colour diagram in the infrared showing those known pre-main sequence eclipsing binaries (∇) and those from this survey (+). The magnitudes of the known pre-main sequence stars are taken from Cargile et al. (2008); Covino et al. (2004); Hebb et al. (2006); Irwin et al. (2007a); Stassun et al. (2004) and Stassun et al. (2007). The dashed lines represent limits on candidature of pre-main sequence binaries (J - H = 0.6 mag, H - K = 0.078 mag) Also plotted are the intrinsic stellar loci of giant stars (••) and dwarf stars (•−) from Bessell and Brett (1988).

ID	Subset	2MASS ID	Right Ascension (J2000.0)	Declination (J2000.0)	Calibrated Magnitudes R	B-R	R-I	Period (days)	Type
UCNS-V-827	5B	16173776-5339210	16:17:37.78	-53:39:20.1204	16.20	2.10	0.31	0.5727	EA
UCNS-V-887	5E	16174598-5332177	16:17:45.99	-53:32:17.4920	16.27	1.60	1.01	2.2913	EA

Table 6.12: **Possible pre-main sequence eclipsing binaries** Table of eclipsing binaries which may be pre-main sequence stars.

6.4.5 Other stars

In cataloguing the variable stars a number of stars have been identified as miscellaneous pulsating stars. These stars are un-categorized either because the lightcurve showed no particular morphology that could be clearly identified or secondary information that was not available was needed. A number of these stars had lightcurves suggesting RR Lyrae type stars but further investigation of their position on the colour-magnitude diagram found them squarely on the main sequence. Examination of some of these stars found secondary long periods, one of these is seen in Figure 6.27.

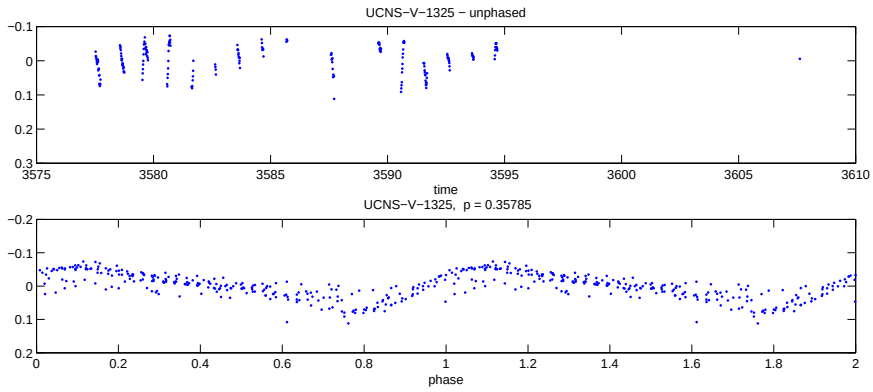


Figure 6.27: Lightcurve of possible secondary modulation of un-categorized pulsating star with top panel showing un-phased data and bottom panel showing phased data.

6.5 Further work

As discussed earlier there are a number of stars in the catalogue that require further observations to truly obtain the most information from them. Further photometric observations of the short-period contact binary stars would be useful to confirm the period and nature of these stars. They are identified as contact binaries but due to the noisy lightcurves this is not yet confirmed. Also needing further followup are the low mass-ratio contact binaries to confirm their nature. Those bright enough low-

mass component candidates require spectroscopic measurements to test the nature of the stars to determine whether there are low-mass components. The Agol-type eclipsing binaries require deeper infra-red measurements in order to compare those with known pre-main sequence stars and identify further pre-main sequence candidates. Already there are two candidate pre-main sequence stars (with infra-red observations from 2MASS) which require spectroscopic followup to find pre-main sequence signatures.

Chapter 7

Conclusions

7.1 Summary

This work describes the observations and results found from a photometric survey of a 0.5° by 0.5° area of the Galactic Plane performed using the 2.2 metre ESO telescope at La Silla, Chile. The dataset comprises a total of 267 images with 204 from a 16 day observation run in 2005 and 63 from a six week observation run in 2002.

Implementation of a new image subtraction algorithm resulted in more than 500,000 lightcurves. Analysis of these for variable stars and transiting extrasolar planet candidates has discovered a large number of variable stars and several good extrasolar planetary transit candidates.

A catalogue of 1475 variable stars has been compiled. There are 1109 eclipsing binaries: 588 contact binaries, 336 semi-detached binaries and 143 detached binaries. Also detected are a large number of miscellaneous pulsating variables.

Examination of the contact binaries have found three which have short periods close to the known period cut off, including two with shorter periods than any previously published. A number of the contact binaries have been found with lightcurve morphologies and periods similar to that of the known low mass-ratio

contact binaries as well as some contact binaries with the same lightcurve shape but with periods outside the known range of low mass-ratio binaries.

Short period contact and detached binaries where the detached binaries show no variation outside the eclipse indicate possible low-mass components. This survey has found 23 contact binaries with periods less than 0.25 days and nine detached binaries with periods less than 1.6 days and appropriate lightcurves. One of these variable stars was found to be bluer than the main sequence of the colour-magnitude diagram indicating that it may contain a white dwarf component.

There are six published eclipsing binaries which are known to have pre-main sequence components. By plotting these binaries and the detached binaries from this survey on a colour-colour diagram with magnitudes from 2MASS we have found two detached binaries with similar characteristics.

Calculation of the binary fraction of the field was made from the observed contact binaries. The binary fraction was calculated to be 66% $\pm$ 4% if the field is of age 10 Gyr and 100% $\pm$ 4% if the field is only 1 Gyr. The most likely age of the field is approximately solar age and using 5 Gyr we find the binary fraction to be 46% $\pm$ 4% which is slightly less than two times larger than that found in solar-type stars.

A search for transiting extrasolar planets was performed and a preliminary list of 31 candidates was found which reduced to 23 when the lightcurves were re-examined. Examination of the images of these candidates removed another nine from consideration.The remaining candidates were re-reduced and examined further. Following this examination we have 3 candidates remaining.

Candidate UCNS-TR-4 has a period of 1.9244 days with a transit depth of 82 ± 9 mmags. Candidate UCNS-TR-7 has a period of 1.4121 with a transit depth of ~ 25 mmags. Candidate UCNS-TR-13 has a period of 1.2016 with a transit depth of 36 ± 2 mmags. The transit signal of this candidate only appears in a single cycle.

The predicted yield of transiting extrasolar planets was nine or ten. In this

research we have detected 3 candidates though they are all of depths at the limit of extrasolar planetary transits. The disparity to the initial predicted yield is due to the lack of observations as compared to the proposed dataset and also due to the uneven time spacing of such. The number of good candidates tallies well with the reduced prediction found from the detection efficiency.

7.2 Suggestions for further study

Following the results found in this research there are several areas requiring further work to make the best use of the data.

7.2.1 Further observations

Further photometric observations with improved time spacing would be of use to improve the periods of known variable stars. In particular period improvements (and confirmations) of the short period contact binaries would help improve the statistics of contact binaries around the period cut-off.

Closely spaced targeted photometric measurements, particularly in multiple filters, would help with determining the nature of the transit candidates. The shape of the transit would be definable and the parameters of the system calculable. Multiple colours allow determination of the shape of the transit as it changes with colour. The well produced by an eclipsing binary will undergo far larger changes than that of a transit. Due to the faintness of the candidates spectroscopic observations on large telescopes are required to obtain radial velocity measurements to a precision to detect planets. Medium resolution measurements would be suitable to determine a maximum mass limit of the companion. Medium resolution measurements are more feasible in the short term particularly following further photometric measurements to refine the nature of the candidates.

Deep infrared observations would be useful for comparison of the detached

eclipsing binaries found in this survey to those known to have pre-main sequence components. The majority of the binaries in this survey are were not observed by the 2MASS survey and thus have no comparable magnitudes. Spectroscopic measurements are also desired of the pre-main sequence candidates found in this work to look for pre-main sequence markers in the spectrum (e.g. lithium absorption (Irwin et al., 2007a)) and would be desired of any further pre-main sequence binaries identified in the infra-red observations.

Further information can be gained by making spectroscopic measurements of those bright low-mass targets to determine the luminosity class and mass of the binary components. Also desired are spectroscopic measurements to further characterize variable star types in particular those variables which have been identified only as pulsating.

7.2.2 Improved reduction algorithm

With improvements made to the reduction algorithm since the full dataset was analysed ideally the full dataset would be re-reduced again. The latest version of the algorithm was used in the manner of microlensing candidates to obtain new photometry of the transit candidates. The script running the algorithm would need some modifications and testing to work on the full sub-image (in general it works on a small section around the candidate star) and a looping function inserted to run the photometry on every star.

7.2.3 SysRem

In this research we have implemented the trend filtering algorithm (TFA) as developed by Kovács et al. (2005) in order to minimise red noise from the lightcurves. Also discussed was the SysRem algorithm developed by Tamuz et al. (2005), this was not implemented. Future work - in particular if a complete re-reduction of the

dataset is carried out should involve implementing the SysRem algorithm. The SysRem algorithm has two particular advantages over the TFA: a determination of the sources of systematics can be made; unlike the TFA SysRem does not increase the white noise in the lightcurves through the removal of red noise. The implementation of SysRem is more complex than that of the TFA and so has not be attempted thus far.